

Result No.	Score	Query %		Length	DB	ID	Description
		Match	Mismatch				
1	1344	100.0	2565	6	AX063459	Sequence	
C 2	129.6	9.6	7218	6	I66494	Sequence 14	
C 3	104	7.7	82781	9	AC103691	AC103691 Homo sapi	
C 4	103.4	7.7	188507	10	AC115738	AC115738 Mus muscu	
5	103	7.7	10732	6	E32986	Gene encodi	
6	102.2	7.6	265985	2	AC087226	AC087226 Mus muscu	
7	101.8	7.6	267182	2	AC116720	AC116720 Mus muscu	
C 8	100.8	7.5	189129	2	AL691432	AL691432 Homo sapi	
9	100.2	7.5	110000	2	AC103465	Continuation (4 of	
10	99.6	7.4	132080	9	AC069368	Homo sapi	
11	98.2	7.3	303091	2	AC084799	AC084799 Mus muscu	
12	97.8	7.3	196204	9	AC073957	AC073957 Homo sapi	
13	97.6	7.3	130540	2	AC079417	AC079417 Mus muscu	
C 14	97.4	7.2	155840	2	AC006495	AC006495 Drosophila	
C 15	97.2	7.2	177277	10	AC147269	AC147269 Mus muscu	
16	97.2	7.2	182065	10	AC138529	AC138529 Mus muscu	
C 17	96.8	7.2	216370	10	AC091332	AC091332 Mus muscu	
18	96.6	7.2	216501	10	AC017795	AC017795 Mus muscu	
19	96.4	7.2	208161	2	AC074145	AC074145 Mus muscu	

[illegible][illegible]

TITLE
JOURNAL
REFERENCE
AUTHORS

Homo sapiens chromosome 15, clone RP11-959H21
 2 (bases 1 to 82781)
 Unpublished

Birren, B., Linton, L., Nuebaum, C., Lander, E., Ali, A., Allen, N.,
 Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhalter, B.,
 Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,
 Chopel, J., Collangelo, M., Collins, S., Collamore, A., Cook, A.,
 Cooke, P., DeAngelis, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S.,
 Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
 Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
 Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
 Jones, C., Katat, A., Karakas, A., Kellis, C., LaRoque, K.,
 Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,
 MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,
 McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrum, J.,
 Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,
 Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
 Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
 Raymond, C., Retta, R., Ribick, M., Riley, R., Rise, C., Rogov, P.,
 Roman, J., Rosatti, M., Roy, A., Santos, R., Schauer, S., Schupack, R.,
 Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
 Topham, K., Travers, M., Travis, N., Triggillo, J., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
 Zainoun, J., Zembek, I., Zimmer, A. and Zody, M.

Only the first 82.8 kilobases of this clone are being submitted. The remainder overlaps accession number AC013553 [WICGR project L2203].

FEATURES
source

	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="15"
	/map="15"
	/clone="RP11-959H21"
repeat_region	clone lib="RPIC1-11 Human Male BAC"
	complement(1..79)
repeat_region	/rpt_family="AluJ/FLAM"
	complement(310..436)
repeat_region	/rpt_family="L2"
	437..883
repeat_region	/rpt_family="L1PA9"
	complement(884..1212)
repeat_region	/rpt_family="L2"
	2288..2502
repeat_region	/rpt_family="MIR"
	complement(2504..2809)
repeat_region	/rpt_family="AluSx"
	complement(2894..3005)
repeat_region	/rpt_family="MIR"
	3273..3312
repeat_region	/rpt_family="AT_rich"
	3408..3576
repeat_region	/rpt_family="FRAM"
	complement(3832..3946)
repeat_region	/rpt_family="MER81"
	4174..4465
repeat_region	/rpt_family="AluSg"
	4572..4689
repeat_region	/rpt_family="MIR"
	complement(5100..5428)
repeat_region	/rpt_family="AluSx"
	5559..5857
repeat_region	/rpt_family="AluJo"
	5861..6005
repeat_region	/rpt_family="AluSx"
	6006..6306
repeat_region	/rpt_family="AluY"
	6307..6469
repeat_region	/rpt_family="AluSx"
	6471..6529
repeat_region	/rpt_family="L1ME"
	6551..6661
repeat_region	/rpt_family="MIR"
	7048..7068
repeat_region	/rpt_family="AT_rich"
	7170..7464
repeat_region	/rpt_family="AluJo"
	7471..7570
repeat_region	/rpt_family="L1MC3"
	7578..7763
repeat_region	/rpt_family="L1MC3"
	7752..8058
repeat_region	/rpt_family="L1MC3"
	8059..8364
repeat_region	/rpt_family="AluSg"
	8365..8450
repeat_region	/rpt_family="L1MC3"
	8450..9281
repeat_region	/rpt_family="L1MC3"
	9282..9574
repeat_region	/rpt_family="AluSx"
	9575..9747
repeat_region	/rpt_family="L1MC3"
	9750..9931
repeat_region	/rpt_family="AluJb"
	9954..10248
repeat_region	/rpt_family="AluJo"
	10307..10937
repeat_region	/rpt_family="L1ME4A"
	11003..11054
repeat_region	/rpt_family="TAGA)n"

[illegible]

Faró, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hago, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Lakocque, K., Lamazares, R., Landers, T., Leshocky, J., Levine, R., Liu, G., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, J., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (22-MAR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 188507)

Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., DeArelano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faró, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hago, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, X., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (22-OCT-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 188507)

Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., DeArelano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faró, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hago, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, X., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (23-DEC-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Dec 23, 2003 this sequence version replaced gi:37806517.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L22816
Center clone name: 474_P_24

FEATURES
source
Location/Qualifiers
1. .188507
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="3"
/map="3"
/clone="RP23-474P24"
/clone_lib="RPCI-23 Female Mouse BAC"
1. .11149
/note="wgs end extension
clone_end:SP6"
complement(1. .396)
/rpt_family="L1"
complement(430. .1594)
/rpt_family="Lx9"
1649. .1744
/rpt_family="TA)n"
1905. .2073
/rpt_family="L1"
2189. .2847
/rpt_family="Lx"
complement(4661. .4821)
/rpt_family="MIR"
4833. .4980
/rpt_family="B1_MM"
complement(5077. .5164)
/rpt_family="MIR"
5206. .5227
/rpt_family="AT_rich"
7500. .7535
/rpt_family="ATGGTC)n"
8932. .9028
/rpt_family="TG)n"
complement(9055. .9331)
/rpt_family="B4A"
9714. .9743
/rpt_family="(TTTG)n"
complement(9857. .10851)
/rpt_family="Lx9"
10858. .10887
/rpt_family="(TC)n"
11150. .11155
/note="clone boundary
clone_end:SP6
site:ECORI"
11274. .11299
/rpt_family="(A)n"
12002. .12056
/rpt_family="(TCTA)n"
complement(12805. .12904)
/rpt_family="PBID10"
13168. .13203
/rpt_family="(TC)n"
13205. .13239
/rpt_family="(TG)n"
13242. .13261
/rpt_family="(GA)n"
15794. .15865
/rpt_family="(TGGA)n"
16699. .16747
/rpt_family="AT_rich"
complement(17093. .17202)
/rpt_family="PBID9"
complement(17224. .17358)

TITLE
JOURNAL
REFERENCE
AUTHORS
COMMENT

```

/rpt_family="B1_MM"
17362. .17395
/rpt_family="(TTTG)n"
complement(17404. .17491)
/rpt_family="B1_MM"
complement(17415. .17516)
/rpt_family="B4A"
17589. .17724
/rpt_family="B1_MM"
18097. .18215
/rpt_family="(TA)n"
complement(18216. .18300)
/rpt_family="L1MC4"
18327. .18392
/rpt_family="(TA)n"
18400. .18430
/rpt_family="(TCTA)n"
19720. .19742
/rpt_family="(A)n"
19839. .19865
/rpt_family="(TTTA)n"
21970. .22017
/rpt_family="AT_rich"
22265. .22334
/rpt_family="CT-rich"
complement(22968. .23114)
/rpt_family="B1_MM"
23936. .24134
/rpt_family="MLT1A"
24437. .24681
/rpt_family="B4A"
24715. .24755
/rpt_family="(CA)n"
24755. .24825
/rpt_family="(GA)n"

```

[illegible]

[illegible]

* 2390	2489: gap of unknown length	* 47281	48753: contig of 1473 bp in length
* 2490	3686: contig of 1197 bp in length	* 48754	48853: gap of unknown length
* 3687	3786: gap of unknown length	* 48854	50036: contig of 1183 bp in length
* 3787	4936: contig of 1210 bp in length	* 50037	50136: gap of unknown length
* 4937	5096: gap of unknown length	* 50137	51270: contig of 1134 bp in length
* 5097	6206: contig of 1110 bp in length	* 51271	51370: gap of unknown length
* 6207	6306: gap of unknown length	* 51371	52534: contig of 1164 bp in length
* 6307	7431: contig of 1125 bp in length	* 52535	52634: gap of unknown length
* 7432	7531: gap of unknown length	* 52635	53751: contig of 1117 bp in length
* 7532	8614: contig of 1083 bp in length	* 53752	53851: gap of unknown length
* 8615	8714: gap of unknown length	* 53852	54995: contig of 1144 bp in length
* 8715	9848: contig of 1134 bp in length	* 54996	55095: gap of unknown length
* 9849	9948: gap of unknown length	* 55096	56243: contig of 1148 bp in length
* 9949	11138: contig of 1190 bp in length	* 56244	56343: gap of unknown length
* 11139	11238: gap of unknown length	* 56344	57370: contig of 1027 bp in length
* 11239	12302: contig of 1064 bp in length	* 57371	57470: gap of unknown length
* 12303	12402: gap of unknown length	* 57471	58713: contig of 1243 bp in length
* 12403	13537: contig of 1135 bp in length	* 58714	58813: gap of unknown length
* 13538	13637: gap of unknown length	* 58814	59843: contig of 1030 bp in length
* 13638	14822: contig of 1185 bp in length	* 59844	59943: gap of unknown length
* 14823	14922: gap of unknown length	* 59944	61019: contig of 1076 bp in length
* 14923	16057: contig of 1135 bp in length	* 61020	61119: gap of unknown length
* 16058	16157: gap of unknown length	* 61120	62326: contig of 1207 bp in length
* 16158	17331: contig of 1174 bp in length	* 62327	62426: gap of unknown length
* 17332	17431: gap of unknown length	* 62427	63761: contig of 1335 bp in length
* 17432	18574: contig of 1143 bp in length	* 63762	63861: gap of unknown length
* 18575	18674: gap of unknown length	* 63862	64972: contig of 1111 bp in length
* 18675	19805: contig of 1131 bp in length	* 64973	65072: gap of unknown length
* 19806	19905: gap of unknown length	* 65073	66224: contig of 1152 bp in length
* 19906	21031: contig of 1126 bp in length	* 66225	66324: gap of unknown length
* 21032	21131: gap of unknown length	* 66325	67394: contig of 1070 bp in length
* 21132	22252: contig of 1121 bp in length	* 67395	67494: gap of unknown length
* 22253	22352: gap of unknown length	* 67495	68559: contig of 1065 bp in length
* 22353	23500: contig of 1148 bp in length	* 68560	68659: gap of unknown length
* 23501	23600: gap of unknown length	* 68660	70794: contig of 2135 bp in length
* 23601	24737: contig of 1137 bp in length	* 70795	70894: gap of unknown length
* 24738	24837: gap of unknown length	* 70895	73758: contig of 2864 bp in length
* 24838	26009: contig of 1172 bp in length	* 73759	73858: gap of unknown length
* 26010	26109: gap of unknown length	* 73859	75386: contig of 1528 bp in length
* 26110	27231: contig of 1122 bp in length	* 75387	75486: gap of unknown length
* 27232	27331: gap of unknown length	* 75487	78567: contig of 3081 bp in length
* 27332	28487: contig of 1156 bp in length	* 78568	78667: gap of unknown length
* 28488	28587: gap of unknown length	* 78668	83211: contig of 4544 bp in length
* 28588	29752: contig of 1165 bp in length	* 83212	83211: gap of unknown length
* 29753	29852: gap of unknown length	* 83312	88000: contig of 4689 bp in length
* 29853	30944: contig of 1092 bp in length	* 88001	88100: gap of unknown length
* 30945	31044: gap of unknown length	* 88101	94340: contig of 6240 bp in length
* 31045	32225: contig of 1181 bp in length	* 94341	94440: gap of unknown length
* 32226	32325: gap of unknown length	* 94441	105003: contig of 10563 bp in length
* 32326	33466: contig of 1141 bp in length	* 105004	105103: gap of unknown length
* 33467	33566: gap of unknown length	* 105104	112336: contig of 7233 bp in length
* 33567	34645: contig of 1079 bp in length	* 112337	112436: gap of unknown length
* 34646	34745: gap of unknown length	* 112437	126183: contig of 13747 bp in length
* 34746	35814: contig of 1069 bp in length	* 126184	126283: gap of unknown length
* 35815	35914: gap of unknown length	* 126284	136888: contig of 10605 bp in length
* 35915	37065: contig of 1151 bp in length	* 136889	136988: gap of unknown length
* 37066	37165: gap of unknown length	* 136989	149641: contig of 12653 bp in length
* 37166	38395: contig of 1230 bp in length	* 149642	149741: gap of unknown length
* 38396	38495: gap of unknown length	* 149742	168565: contig of 18824 bp in length
* 38496	39843: contig of 1348 bp in length	* 168566	168665: gap of unknown length
* 39844	39943: gap of unknown length	* 168666	184976: contig of 16311 bp in length
* 39944	41035: contig of 1092 bp in length	* 184977	185076: gap of unknown length
* 41036	41135: gap of unknown length	* 185077	203645: contig of 18569 bp in length
* 41136	42239: contig of 1104 bp in length	* 203646	229531: contig of 25786 bp in length
* 42240	42339: gap of unknown length	* 229532	229631: gap of unknown length
* 42340	43436: contig of 1097 bp in length	* 229632	265985: contig of 36354 bp in length.
* 43437	43536: gap of unknown length	Location/Qualifiers	
* 43537	44696: contig of 1160 bp in length	1. .265985	
* 44697	44796: gap of unknown length	/organism="Mus musculus"	
* 44797	45930: contig of 1134 bp in length	/mol_type="genomic DNA"	
* 45931	46030: gap of unknown length	/db_xref="taxon:10090"	
* 46031	47180: contig of 1150 bp in length	/chromosome="16"	
* 47181	47280: gap of unknown length	FEATURES	
		source	


```

/clone="RP23-103M13"
/clone_lib="RPCI mouse BAC library 23"

ORIGIN
Query Match          7.6%; Score 102.2; DB 2; Length 265985;
Best Local Similarity 46.8%; Pred No. 9.7e-12;
Matches 374; Conservative 0; Mismatches 422; Indels 4; Gaps 2;

QY      536 AAGAAGAAACCAAGAGCGCCACAGGCGCAGGCCAAGGAAGCGGAGGAGGCGGAGGAG 595
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      62759 AAGAGAGAAAGAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAA 62818

QY      596 AAAGAACCGCGCAACAACGGCCAGAGCAGAAAGCAGAGGCAGAAAGAGAGAGAG 655
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      62819 AAAAAGAAAAAGAGAGAGAAAAAGAAAAAGGAGGAGGAGGAGGAGGAGGAG 62878

QY      656 CGAAGACCGCGAGAGCCCAAAAGCCCAACCAAGAAAGACAGCGGAGAGAGAGAG 715
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      62879 AGAGGAAAAAGAAAAAGAAAAAGAGGAGGAGAAAAAGAAAAAGAGAGAGAG 62938

QY      716 ACCACAGCGCGAGACCGAGCGCCCGGCGAGAGGAGGAGAACACCCACCGCGAG 775
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      62939 AAAAAAGAAAAAGAA---GAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA 62995

QY      776 CAAGAAAGACCACCCCAACACAGAGAGAGGAGGACACAGAGAGCCAAAGAGAGAG 835
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      62996 AAGAGGAGAGAAAAAGAAAAAGAAAAAGGAGAGAGAAAAAGAAAAAGGAGAGG 63055

QY      836 ACAACAGCGGAGAGCGGAGGAGCACCGCGCGCCACAGAGAGACAGAGAAAG 895
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      63056 AAAAAAGAAAAAGAAAAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 63115

QY      896 ACCAGAAAAAGGAGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 955
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      63116 AAAAAAGAGAGAGAGG-GGAGAAAAAGAAAAAGAAAAAGAAAAAGAGAGAGG 63174

QY      956 GCCAGACAAGCCGAAAAAGAGAGCGGCGAAACGGCGCAACACAGAACAGAGC 1015
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      63175 NGNNGAAAAAGAGAGAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 63234

QY      1016 ACGGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1075
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      63235 GAAGAAAAAGAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 63294

QY      1076 ACCAGCGCCACCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1135
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      63295 AAAAAAGAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 63354

QY      1136 AAAGCAGACCCCAACGAAAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1195
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      63355 AGAAAAAGAAAAAGAAAAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 63414

QY      1196 GCAGAGAAAAAGGCCCCAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1255
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      63415 AGAAAAAGAAAAAGAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 63474

QY      1256 CGGACGAGGCCCCAACACCAAGCCCAAGAGACAGAGGAGGAGGAGGAGGAGGAG 1315
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      63475 GGCGGGGGGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAGAGAGGAG 63534

QY      1316 CACACAGCAGAGAGAGAGA 1335
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      63535 AGAGAGGAGAGAGAGAGAGA 63554

RESULT 7
LOCUS       AC116720
DEFINITION Mus musculus clone RP23-223N19, WORKING DRAFT SEQUENCE, 19
            unnumbered pieces.
ACCESSION   AC116720
VERSION     AC116720.4 GI:30018035
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
```


QY	62	AGCAAAACCCAGGGACACAGCGGACCGAAACAGAAAGAGAGAGGGGGCCAAACCGGGCCG	121
Db	170246	AGACACAGACAGACACAGAGAGACGAAACAGGGAGAGACAGAGAGACAGAGAGAGAG	170187
QY	122	CCGACAGAGAAACACAGCCCAAGCAGGGGACCGCGCCAGGAGCAACCCAGG	181
Db	170186	TGACAGAGAGAGGAGAGAGAGAGAGAGAGAGACAGACACAGAGAGACAGACAGGGAG	170127
QY	182	CCGCGGGCACACAAACCGGGGACAGAGGGCCCGGACAAACCGAGGAGGAGAGACACC	241
Db	170126	AAACAGAGAGACAGAGACCGAGAGACAGAGACAGAGAGAGGACAGACAGACAGAGAG	170067
QY	242	CCGCAAAACCCCGGAAACAGCCGAGAGCCCGAGCCGAGCAGGAGCCAGCCAAAGAA	301
Db	170066	AGAGACACACAGACAGACACAGAGAGAGCAGAACCGGGGAGAGACAGAGAGACAGAGAA	170007
QY	302	AGCAACAGAAACGGCCCGACACAGCGGACCAACAGAGACAGAAAGCCCGAGCGCCCCGAGA	361
Db	170006	AC--AGACACAGACAGAGAGGACAGACAGAGAGAGACAGACAGAGAGCAGAACAGG	169944
QY	362	GCAGAGAGACGCACACACACACAGAGGGCCCGGCGAGCAGAGAACCGAGAACCCAGGG	421
Db	169948	AGAGACAAAGACACAGAGAGAGAGAGACACAGAGAGAGAGATAGAGAGAGGCGAC	169889
QY	422	CAGGACAAACCCGAGGACCGCCAGGAGACAGGAGAAAGCACCAGGAGACAGAGAGA	481
Db	169888	AGAGACAGAGACACAGACACAGAGCGAACHAGGGAGAGACAGAGAG-----AG	169835
QY	482	GCAGCGAGAGGACACAGAGAGAAAGCCCCACAGCGCCCAACCCACCGCGGAGAGAG	541
Db	169834	AGACAGAGAGAGGCGACACAGAGAGAGAGAGACACACAGAGAGACAGAACAGGGAG	169775
QY	542	AAACCAAGAAAGGCCACAGGGCAGCCAGGAAGCCAGAAAGGCCGAGAGGAGAGAA	601
Db	169774	AAACAGAGACAGAGAGCGAGAGAGACACACAGAGAGAGGCGACAGACAGACAGAGAG	169715
QY	602	CCCGCAACACACGGGCAGAGACAAAGACAGAGCCCAAGAGCAGCAAAAGACCGAAGA	661
Db	169714	AGCAAAACAGACACACAGAGAGACAGACAGAGCCAGACCGGGAGACAGAGACAGAGAGAA	169655
QY	662	CCGCCAGAGCCCCAAAGCCCAACCAAGAAAGACCGGACGAGAGGAAACCCCAACCAAC	721
Db	169654	AGAGACAGAGACAGAGAGAGGCGACAGACAGAGACACACAGACAGAGAGAGAGAAC	169595
QY	722	AGCGCAGACCGGAGGACCGCGGCGGAGAGGAGAACACCCACACCCCGAGGCCACAAA	781
Db	169594	AGGGAGAGACAGAG-----AGACACAGACACAGAGAGGCGCAACACAGACAGAGACAGA	169540
QY	782	GACCAACCCCAACACAGAGGAGAGAGAAACACACAGAGAGCCCAAGGAGAGGACAA	841
Db	169539	CAGAGAGACAGATGGGGAGAGACACAGAGACAGAGACAGAGACAGAGAGAGGCGAGAC	169480
QY	842	CGGCGACAGCGGGAGCAGCCCGCGCACAGAAAGACAGAGAAACAGCAGAGAGACCCAG	901
Db	169479	GAGACAGAGAGAGACAGACACACACAGACAGACAGAGCAGAAACAGGGAGAGACG----	169424
QY	902	AAAGGGACCGCAGGGAGAAAGCCACAGAAAGGAAAGCCACCGCGGAGGAAAGCCAG	961
Db	169423	GAGACAGAGAGGCGACAGACAGACAGAGAGAGAGACAGACACAGAGCAGACAGGGAGA	169364
QY	962	CAAGCCGAAAGAGCAGGCGAAACCGGGCGAAACACAGAGACAAAGGAGCCGAAAAACCC	1020
Db	169363	GACAGAGAGAGAGACAGAGAGAGGCGACAGAGAGAGAGAGAGAGAGACAGACAGAG	169304
QY	1021	AGAGAAAGGGAGAGAAACACAAACACCCAGGGGGGCAAAAGGGAGCCGAAAAACCC	1080
Db	169303	GAGAGCGAAACAGGGAGAAACAGAGACAGAGCGAGAGACAGAGCGGAGAGAGAGAGAG	169244
QY	1081	GCCCAACCCGAGAGGACAGCCCAAGCACAACAAAGGCGCAACAAAGGGGAAAGGCCAAA	1140
Db	169243	GACAGACACAGAGAGACAGACAGACAGAGAGACAGACAGAGAGACAGAACCGGGAGAG	169184
QY	1141	CACACCCCAACGAAAGACAGGGACAGAAAGCGCGCGGAGAGAAAGAAAGAGGCGCAGA	1200

[illegible]

```

Db      17316 CACGAGAGCAGAAAAAGGTGAAGCAGAAAAATAAGCAGAACAGCAGAGCAGAAAGGTGA 17375
QY      984 AACGGCGGCAACACAGAACAAAGACGCGAGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1043
Db      17376 AGCAAGAAACACAGACAGAAAGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 17435
QY      1044 ACAACCCAGGGGGGCAAAAGGAGCGAAGAAACCCAGCCACCCAGAGAGAGAGAGAGAGAG 1103
Db      17436 GAATCAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 17495
QY      1104 AAGCAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1159
Db      17496 AAGCAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 17555
QY      1160 AGGCGACAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1219
Db      17556 AGAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 17615
QY      1220 AGGCGCAGCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1254
Db      17616 GAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 17650

```

```

RESULT 10
AC069368
LOCUS      132080 bp      DNA      linear      PRI 15-MAY-2002
DEFINITION Homo sapiens chromosome 15, clone CTD-2017F17, complete sequence.
ACCESSION AC069368
VERSION   AC069368.14  GI:20800366
KEYWORDS  HTG.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 132080)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 15, clone CTD-2017F17
Unpublished
2 (bases 1 to 132080)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
Boguslavskiy,L., Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collimore,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domingo,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levine,R., Liu,C., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheters,R.,
Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mienga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neil,D., Olivier,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Testfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (28-MAY-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 132080)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collimore,A.,
Cooke,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gort,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS

```

```

Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,
MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,
Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
Olivier,T., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
Orliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Testfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (20-APR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 132080)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collimore,A.,
Cooke,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gort,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,
MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,
Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
Olivier,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Riley,R., Rise,C., Rogov,P.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Testfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (15-MAY-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 15, 2002 this sequence version replaced gi:20219280.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RW/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L8893
Center clone name: 201_F_17
-----
Location/Qualifiers
1. .132080
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosomes="15"
/map="15"
/clone_lib="CTD-2017F17"
/clone_lib="CTD1 Human BAC"
438. .602
/rpt_family="MER5B"
615. .663
/rpt_family="L3"
1348. .1492
/rpt_family="FLAM_C"
complement(1670..2021)
/rpt_family="L1MB4"

```

```

TITLE
JOURNAL
COMMENT

```

```

FEATURES
source
repeat_region
repeat_region
repeat_region
repeat_region

```



```

Db      16821 GAAGAAGAACTAGAAGA 16838
RESULT 11
AC084799
LOCUS   AC084799
DEFINITION AC084799 303091 bp DNA linear HTG 17-NOV-2000
          Mus musculus chromosome 16 clone RP23-197M9, WORKING DRAFT
          SEQUENCE, 101 unordered pieces.
ACCESSION AC084799
VERSION   AC084799.1 GI:11192127
KEYWORDS  HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE    Mus musculus (house mouse)
ORGANISM  Mus musculus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 303091)
AUTHORS   DOE Joint Genome Institute.
TITLE     Sequencing of Human Chromosome 16
JOURNAL   Unpublished
REFERENCE 2 (bases 1 to 303091)
AUTHORS   DOE Joint Genome Institute.
TITLE     Direct Submission
JOURNAL   Submitted (17-NOV-2000) Production Sequencing Facility, DOE Joint
          Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT   -----Genome Center
          Center: Joint Genome Institute
          Center Code: JGI
          Web site: http://www.jgi.doe.gov
          -----
Project Information
Center Project Name: 0
Center clone name: RPCI-23_197M9
-----
Summary Statistics
Consensus quality: 152568 bases at least Q40
Consensus quality: 175579 bases at least Q30
Consensus quality: 186949 bases at least Q20
Estimated insert size: 198300; agarose-fp estimation
Estimated insert size: 293091; sum-of-contigs estimation
Quality coverage: 4.85 in Q20 bases; agarose-fp estimation
Quality coverage: 3.28 in Q20 bases; sum-of-contigs estimation
* NOTE: This is a 'working draft' sequence. It currently
* consists of 101 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
1117 1116: contig of 1116 bp in length
1217 1216: gap of unknown length
2337 2336: contig of 1320 bp in length
2637 2636: gap of unknown length
4058 4157: gap of 1421 bp in length
4158 5245: contig of 1088 bp in length
5246 5345: gap of unknown length
5346 6461: contig of 1116 bp in length
6462 6561: gap of unknown length
6562 7725: contig of 1164 bp in length
7726 7825: gap of unknown length
9144 9243: contig of 1318 bp in length
9244 10492: contig of 1249 bp in length
10493 10592: gap of unknown length
11667 11766: gap of 1074 bp in length
11767 12950: contig of 1184 bp in length
12951 13050: gap of unknown length
13051 14285: contig of 1235 bp in length
14286 14386: gap of unknown length
14387 15478: contig of 1093 bp in length
15479 15578: gap of unknown length
15579 16779: contig of 1201 bp in length
16780 16879: gap of unknown length
16880 18278: contig of 1399 bp in length
18279 18378: gap of unknown length
18379 19452: contig of 1074 bp in length
19453 19552: gap of unknown length
19553 21012: contig of 1460 bp in length
21013 21112: gap of unknown length
21113 22169: contig of 1057 bp in length
22170 22269: gap of unknown length
22270 23367: contig of 1098 bp in length
23368 23467: gap of unknown length
23468 24559: contig of 1092 bp in length
24560 24659: gap of unknown length
24660 25925: contig of 1166 bp in length
25926 27036: contig of 1111 bp in length
27037 27136: gap of unknown length
27137 28245: contig of 1109 bp in length
28246 28345: gap of unknown length
28346 29657: contig of 1311 bp in length
29657 30883: gap of unknown length
30884 30983: gap of unknown length
30984 32125: contig of 1142 bp in length
32126 32225: gap of unknown length
32226 33397: contig of 1172 bp in length
33398 33497: gap of unknown length
33498 34645: contig of 1148 bp in length
34646 34745: gap of unknown length
34746 35894: contig of 1149 bp in length
35895 35994: gap of unknown length
35995 37156: contig of 1162 bp in length
37157 37256: gap of unknown length
37257 38396: contig of 1140 bp in length
38397 38496: gap of unknown length
38497 39570: contig of 1074 bp in length
39571 40852: contig of 1182 bp in length
40853 40952: gap of unknown length
40953 42116: contig of 1164 bp in length
42117 42216: gap of unknown length
42217 43372: contig of 1156 bp in length
43373 43472: gap of unknown length
43473 44603: contig of 1131 bp in length
44604 44703: gap of unknown length
44704 45924: contig of 1221 bp in length
45925 46024: gap of unknown length
46025 47174: contig of 1150 bp in length
47175 47274: gap of unknown length
47275 48744: contig of 1470 bp in length
48745 48844: gap of unknown length
48845 49969: contig of 1125 bp in length
49970 50069: gap of unknown length
50070 51135: contig of 1066 bp in length
51136 51235: gap of unknown length
51236 52690: contig of 1455 bp in length
52691 52790: gap of unknown length
52791 53920: contig of 1130 bp in length
53921 54020: gap of unknown length
54021 55101: contig of 1081 bp in length
55102 55201: gap of unknown length
55202 56324: contig of 1123 bp in length
56325 56424: gap of unknown length
56425 57583: contig of 1159 bp in length
57584 57683: gap of unknown length
57684 58822: contig of 1139 bp in length
58823 58922: gap of unknown length
58923 60547: contig of 1625 bp in length
60548 60647: gap of unknown length
60648 61743: contig of 1096 bp in length
61744 61843: gap of unknown length

```

QY		671	CCCAAAAGGCCCAACCTAAGAAGACACGGGACGAGAGAAAACCCATCACAACAGCGCCGAG	94100
Dd		94042	CACNAGAGTAGGCAGACAGCGAACGAAGAAGAGCGA-AGGAAGACCAGAGAGGAGGAAAAACA	94101
QY		731	CCGAGGCACCGGGCGGAGAGAGAACACCCACACCGCCGAGAGCCACAAAGAGCACCAACCCC	94102
Dd		94101	CCGAGACCCAGAAAACAGAGCGCGGAAGACCACACGACGCAACGACACGCGCACACGCGCACACGCG	94103
QY		791	CAAAACAGAGGAGAGGAGAGAACACCCAGGAAGCCCAAGGAAGGAAGAGGACAAACAGCGCGAGAC	850
Dd		94161	CAAAGAGGAGCGGAAGGAAGAGAGAAAGAAAGAAAGAAAGAGAAAGGGAAGGAGAAAG	94220
QY		851	GGGAGGACGCGCGCGCCACAGAAAGACAGAGAAAAGCAGAGAGACCCAGAAAAGGGAC	910
Dd		94221	AGAAAGAGAGAGGAGGAGAGAAAGAGAGAGAAAGGAAGGAAGGAAGAAAGAGAAAGAGAGGC	94280
QY		911	CGAGGGAGAAAGGCCAGCAGNAAGGAAGGACCCGCGAGGAGAAAGCCAGACAAAGCCGAA	970
Dd		94281	AGAGAGAGAGAGAGAAAGAGAGAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGCGAGAGAG	94340
QY		971	AAGAAGCAGGCGMAAACGGGCGCAACACAG-ACAAAGAACCGCAGAGAACCGGAAGAGG	1029
Dd		94341	AACCGAGAGCGGAGGAGAGACACAGAGAGAAACACAGAGCAGAGAGAGAAACAGAGA	94400
QY		1030	GAGGAAGAACACAAACAAACCCAGGGGGGCAAGGHCAGGACCAAAACCCAGCCCCCG	1089
Dd		94401	GAGACGGGAACAGACAG	94460
QY		1090	AGAAGGAGACGCCAGACACAAAAGGGCGCAACAAAGGGGAAGGCGCCAAAAGAGCACACCCCA	1149
Dd		94461	ACAG	94520
QY		1150	ACGAAGAAGACAGGAGACAGAAAAGCGCGCCGAGAAAAGAAAGAGAGCGCCAGAGAAAAGG	1209
Dd		94521	CGAGACAGCAGAGAGGAG	94580
QY		1210	CCCGAGGCGAGGCGCCAGCCAAAGGAGA	1238
Dd		94581	CGCCANGAGACGCCAGAGAGGAGACGAGA	94609
RESULT	12			
AC073957				
LOCUS		196204 bp	DNA	linear PRI 09-JAN-2002
DEFINITION		Homo sapiens BAC clone RP11-449P15 from 7,	complete sequence.	
ACCESSION		AC073957		
VERSION		AC073957.7	GI:17149791	
KEYWORDS		HTG.		
SOURCE		Homo sapiens (human)		
ORGANISM		Homo sapiens		
REFERENCE		Makryiotis, S., Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 196204) Sulston, J.E. and Waterston, R. toward a complete human genome sequence Genome Res. 8 (11), 1097-1108 (1998) 99063792 FUBMED 9847074 2 (bases 1 to 196204) Leonard, S., Cotton, M., Hawkins, M. and Spalding, L. The sequence of Homo sapiens BAC clone Rp11-449P15 Unpublished (2001) 3 (bases 1 to 196204) Waterston, R.H. Direct Submission Submitted (07-JUL-2000) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA 4 (bases 1 to 196204) Waterston, R.H. Direct Submission Submitted (29-NOV-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA		

MO 63108, USA
 5 (bases 1 to 196204)
 Waterston,R.H.
 Direct Submission
 Submitted (03-JAN-2002) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 6 (bases 1 to 196204)
 Waterston,R.
 Direct Submission
 Submitted (09-JAN-2002) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 On Nov 29, 2001 this sequence version replaced gi:14589737.
 ----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu/gsc>
 Contact: sapiens@wustl.wustl.edu
 ----- Summary Statistics

 Center project name: H_NH0449P15

NOTICE: This sequence may not represent the entire insert of this
 clone. It may be shorter because we only sequence overlapping
 clone sections once, or longer because we provide a small overlap
 between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate
 chemistry, or covered by high quality data (i.e., phred quality >=
 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by sequence
 from more than one subclone; and the assembly was confirmed by
 restriction digest.

MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and
 sequencing collaboration between the NHGRI Chromosome 7 Mapping
 Project (Eric D. Green, Director), John D. McPherson in the
 Department of Genetics (Washington University), and the Washington
 University Genome Sequencing Center. For additional information
 about the map position of this sequence, see
<http://www.nhgri.nih.gov/DIR/STB/CHR7>, send
 mail to: egreen@nhgri.nih.gov, or see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male
 donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E.,
 Tatenio,M., Catanese,J.J. and de Jong,P.J. (1998) An improved
 approach for construction of bacterial artificial chromosome
 libraries. Genomics 51:1-8. The clone may be obtained either from
 Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong
 and coworkers at the Roswell Park Cancer Institute
 (<http://bacpac.med.buffalo.edu>)
 VECTOR: pBAC3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-1244M4; the clone sequenced
 to the right is RP13-580H13. Actual start of this clone is at base
 position 1 of RP11-449P15; actual end is at base position 196204 of
 RP11-449P15.

Data from AC091729 was used to finish this clone, AC073957.
 Polymorphisms have been identified between AC073957 and AC091729.

Unresolved simple sequence repeats, base positions 26410 to 27787
 and 63334 to 64953.

FEATURES source

Location/Qualifiers
 1..196204
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="7"
 /map="7"

	/clone="RP11-449P15"	
	1..278	/clone_lib="RPCI-11"
misc_feature	/note="match to EST BG992151 (NID:g14396221)"	
misc_feature	203..598	
misc_feature	/note="match to EST BF814714 (NID:g12147717)"	
	1010..1156	
repeat_region	/note="match to EST F06563 (NID:g672150)"	
	1122..1414	
misc_feature	/rpt_family="Alu"	
	1344..1863	
repeat_region	/note="similar to Homo sapiens EST BF928866 (NID:g12326994)"	
	1932..2224	
misc_feature	/rpt_family="Alu"	
	2189..2393	
repeat_region	/note="match to EST F06563 (NID:g672150)"	
	2270..2636	
repeat_region	/rpt_family="L1"	
	2393..2422	
repeat_region	/rpt_family="(CACCC)n"	
	2745..3367	
misc_feature	/rpt_family="L1"	
	4057..4262	
misc_feature	/note="match to EST BI039739 (NID:g14446365)"	
	5018..5019	
misc_feature	/note="match to EST AW877271 (NID:g8015626)"	
	5068..5455	
misc_feature	/note="match to EST BF087410 (NID:g10893120)"	
	5193..5455	
misc_feature	/note="match to EST AA004851 (NID:g14485598) zh92f02.rl"	
	5405..5761	
misc_feature	/note="match to EST AW877271 (NID:g8015626)"	
	5829..5833	
repeat_region	/note="match to EST BG992151 (NID:g14396221)"	
	5844..6140	
misc_feature	/rpt_family="Alu"	
	6107..6377	
misc_feature	/note="match to EST BE073441 (NID:g8420427)"	
	6136..6426	
misc_feature	/note="match to EST BF874571 (NID:g12264701)"	
	6174..6844	
misc_feature	/note="match to EST AUI33804 (NID:g10994343)"	
	7097..7195	
misc_feature	/note="similar to Mus musculus EST AA021831 (NID:g1485585) mh82c08.rl"	
	7108..7195	
misc_feature	/note="match to EST AI200352 (NID:g3752958) qf92h12.xl"	
	7108..7195	
misc_feature	/note="match to EST BF516594 (NID:g11601773)"	
	7108..7195	
misc_feature	/note="match to EST BG709448 (NID:g13987794)"	
	7108..7195	
misc_feature	/note="match to EST T16043 (NID:g518205)"	
	7109..7195	
misc_feature	/note="match to EST AL527751 (NID:g12791244)"	
	7117..7199	
misc_feature	/note="match to EST AA004851 (NID:g1448598) zh92f02.rl"	
	7117..7195	
misc_feature	/note="match to EST BF087410 (NID:g10893120)"	
	7141..7195	
misc_feature	/note="match to EST AI174672 (NID:g6361050)"	
	7150..7195	
misc_feature	/note="match to EST BG034421 (NID:g12427714)"	
	7628..7711	
misc_feature	/note="similar to Mus musculus EST W65144 (NID:g1372795) me02f03.rl"	
	7630..7711	
misc_feature	/note="match to EST AI174672 (NID:g6361050)"	
	7630..7711	
misc_feature	/note="match to EST AI200352 (NID:g3752958) qf92h12.xl"	
	7630..7711	
misc_feature	/note="match to EST AL527751 (NID:g12791244)"	

[illegible]


```

* 153671 154319: contig of 649 bp in length
* 154320 154399: gap of unknown length
* 154743 154743: contig of 344 bp in length
* 154744 154823: gap of unknown length
* 154824 154922: contig of 599 bp in length
* 155423 155502: gap of unknown length
* 155503 155840: contig of 338 bp in length.
FEATURES
    source      Location/Qualifiers
1..155840
    /organism="Drosophila melanogaster"
    /mol_type="genomic DNA"
    /strain="y; cn bw sp"
    /db_xref="taxon:7227"
    /chromosome="3"
    /map="90E1-90F10"
    /clone="BACR48113 (D522) RPCI-98 48.1.13"
    /clone_lib="RPCI-98 (Roswell Park Cancer Institute
Drosophila melanogaster BAC library, partial EcoRI in
pBACE3.6"
ORIGIN
Query Match      7.2%; Score 97.4; DB 2; Length 155840;
Best Local Similarity 45.5%; Pred. No. 1.1e-10;
Matches 351; Conservative 0; Mismatches 409; Indels 11; Gaps 2;
QY 279 GCAGGAGGAGCGACAAAGAAAGCAACAGAAACCGCCGACCGAGCGGAGCGGAGAG 338
Db 135448 GGAGGACAAAGGGGAAACAGAGAAAGNGAAACAGGGGGGAGGGGGGGGAGCGGNCAG 135389
QY 339 CAGAGACCCCGAGCGCCCC---GAGAGGAGAGAGCGCACACACACAGAGGGCCCGC 395
Db 135388 CAGCAGCAGGGCGCAACANAGAGGAGCAACAGANAGNGGCGACACACGCGGAGCAGNNG 135329
QY 396 GCGAAGCAGGAAACAGAAACACGAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 455
Db 135328 CGGGCGGCGACAAAGCGGGGAGACACGAGAAACGCGAGAGGAGGAGGAGGAGG 135269
QY 456 AGGAGAGCAGCGGGAGCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 515
Db 135268 CNGCGGNNCGACAGAGGAGCAACANACGNCNAGCAGGGGAGGAGGAGGAGGAGG 135209
QY 516 CGCCCAACACACACCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 575
Db 135208 GGAGGAAAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 135149
QY 576 CGAGAAAGGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 627
Db 135148 GGACAAAGGGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 135089
QY 628 AGCAGAGCCAAAGAGAGCAGCAAAAGAGCCGAGAGGAGGAGGAGGAGGAGGAGGAG 687
Db 135088 GGCAAGGCAACAGGCAACCCANAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 135029
QY 688 AGAAGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 747
Db 135028 GGGCAGCCACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 134969
QY 748 GAGGAGAAACACCCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 807
Db 134968 GAGGAGAAACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 134909
QY 808 GAGAACACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 867
Db 134908 CAAAGAAACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 134849
QY 868 CCACAGAGCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 927
Db 134848 AACCGGAGACCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 134789
QY 928 GCAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 987
Db 134788 GAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 134729

```

```

QY 988 GGGCAACACAGAACAAAGAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1038
Db 134728 AAACCCGAGAGCGGAGCAGACACCAAGACAGGAGGAGGAGGAGGAGGAGGAG 134678

RESULT 15
AC147269/c
LOCUS      AC147269      177277 bp      DNA      linear      ROD 15-MAY-2004
DEFINITION Mus musculus BAC clone RP24-281E6 from chromosome 18, complete
            sequence.
ACCESSION  AC147269
VERSION    AC147269.3  GI:45237303
KEYWORDS   HTG.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
REFERENCE  1 (bases 1 to 177277)
            Birkhoff, A., Haglund, K. and Bielicki, L.
            The sequence of Mus musculus BAC clone RP24-281E6
            Unpublished (2001)
REFERENCE  2 (bases 1 to 177277)
            Wilson, R.K.
            Direct Submission
            Submitted (07-NOV-2003) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
REFERENCE  3 (bases 1 to 177277)
            Wilson, R.K.
            Direct Submission
            Submitted (17-DEC-2003) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
REFERENCE  4 (bases 1 to 177277)
            Wilson, R.K.
            Direct Submission
            Submitted (05-MAR-2004) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
REFERENCE  5 (bases 1 to 177277)
            Wilson, R.K.
            Direct Submission
            Submitted (15-MAY-2004) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
            On Mar 6, 2004 this sequence version replaced gi:39979572.
            ----- Genome Center
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site: http://genome.wustl.edu
            Contact: submissions@wustl.wustl.edu
            ----- Summary Statistics
            -----
            Center project name: M_BB0281E06
            -----

```

NOTICE:

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone, fosmid clone or direct clone walk sequence. Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to obtain the consensus sequence; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:

The RPCI-24 BAC library has been constructed by Pieter de Jong and coworkers (<http://www.chori.org>) from male C57BL/6J mouse spleen and/or brain genomic DNA. The clone and detailed information can be

obtained from Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:

NEIGHBORING SEQUENCE INFORMATION:
This sequence is the entire insert of the clone. This clone is overlapped by AC139334.

FEATURES

source

1. .177277

l. : I)Z//
/organism="Mus musculus"

```
/organism="Mus musculus"
/tax_type="genomic DNA"
```

```
/mob_type="germ" /db_xref="taxon:10090"
```

```
/w_xict = c("18")
/chromosome="18"
```

```

/ map="18"

```

```
/clone="RP24-281E6"
```

```
/clone lib="RPCI-24"
```

105840⁻.105910

```
/product="tRNA-Ser"
```

/note="Likely pseudogene (HMM SC=32.06 / sec struct

$$SC = -10.06$$

ORIGIN

Query Match	7.2%;	Score 97.2;	DB 10;	Length 177277;
Best Local Similarity	46.8%;	Pred. No. 1.2e-103;		
Matches	378;	Conservative 0;	Mismatches 423;	Indels 7; Gaps 2;
QY	398	GAACGAGCAACCGAATACCAGGGCAGACACACACCGAAGGAGGACACAGCCAGGAGCAACAG	457	
DB	97498	GAAGGAGCGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG	97439	
QY	458	GAGAAGACACCGGAGCAGACAGAGAGGACCGAGAGGGACCCAGCAGGAAAGGCCCCACGAGCG	517	
DB	97438	CGAAGGAAGGAGGAGGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG	97381	
QY	518	CCACACCCACGACCCCGAAGAAAGAAACCAAGAGGCCACACAGGGCAGCCCAAGGAAGCCG	577	
DB	97380	CAGACAGACAGCGGAAGGAGGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG	97321	
QY	578	AGAAAGGCCCGAGCAGAGAAAGAAACCGCGCAACACCGGCAGACAGACAAAGAACGAGAGCCCA	637	
DB	97320	AAAGAAAGACAGACAGGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	97261	
QY	638	AAAGAAAGCAGCAAAAGACCCGACCGCCAGAGCCCAAGCCCAACCAAGAAAGACAC	697	
DB	97260	GACAGGGAAGGAAGGAAGGAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAG	97201	
QY	698	GGGACGAGAGGAAACCCAAACCAACAGCCGAGAGCCGAGGACCCGGGCCGAGAGGAGAAACA	757	
DB	97200	AGAAAGAAAGAAACAGACAGACAGACAGGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAG	97141	
QY	758	CCACACCCCGGAAGCCCAAAAAGACCAACCCCAACACAGAGGAGGAAGAGAACCAACA	817	
DB	97140	AGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG	97086	
QY	818	GGAAAGCCAAAGGAAGGAGCAACAGCGGGCAGACGGGGAGCAGCCCGGCCCCACAGAGA	877	
DB	97085	GAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG	97026	
QY	878	CAGAGAAAGCCGAGAGACCAAGAAAGGACCCGAGGAGAAAGGCCAGCAGCAAGAAAGGA	937	
DB	97025	AGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG	96966	
QY	938	AAGCAGCCCGGAGGAAGCCAGCAAGCCGGAAGAGCAGGCGAAAGCAGGCGAAACCGGGCGAAC	997	
DB	96965	AAGGAAGGAGGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG	96906	
QY	998	GAACAAAGACCGCAGGAAACGGAAGAGAGAGGGAGGAGGAACACAAAAACAACCCAGGGG	1057	
DB	96905	GAAGGAAGGAGGAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAG	96846	
QY	1058	GGCAAAAGGGACCCGAAACCCAGCCACCCGAGAGGAGAGCAGCCTAAGCAAAAAAGG	1117	
DB	96845	GAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG	96786	
QY	1118	CAACAAAGGGGAAGGCCAAAGCAACCCCAACGAAAGACAGGGACAGAAAGAGCGCGC	1177	

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 20, 2004, 17:04:23 ; Search time 728 Seconds
(without alignments)
9691.239 Million cell updates/sec

Title: US-10-019-341-4_COPY_256_1599
Perfect score: 1344
Sequence: 1 aagacacagcaggacacgc.....gagaagagagaagaccccg 1344

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_23Sep04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1344	100.0	2565	4	Aaf28296
2	103	7.7	10732	3	AAAF20844
3	102	7.6	6755	2	AAV21511
C 4	97	7.2	42999	8	ABX11086
C 5	97	7.2	114955	2	AAAX53491
C 6	94.8	7.1	41907	9	ADA37416
C 7	94.8	7.1	42998	8	ADL4747
C 8	94.8	7.1	42999	6	ABX5032
C 9	94.8	7.1	42999	10	ADL61411
C 10	91.2	6.8	6741	3	AAAF10595
C 11	89	6.6	5059	2	AAAX84332
C 12	88.8	6.6	2318	5	ADL63176
C 13	88.8	6.6	3662	4	ABL05274
C 14	88.6	6.6	1200	6	ABQ39210
C 15	88.6	6.6	1200	6	ABQ39211
C 16	87.6	6.5	1840	10	ADC87550
C 17	84.8	6.3	1762	6	ABQ20939
C 18	84.8	6.3	1762	6	ABQ20938
C 19	83.2	6.2	42533	8	AAAB56114
C 20	83.2	6.2	42533	9	ADA02476
C 21	83.2	6.2	42533	10	ADB72215

C 22	81.8	6.1	7033	3	AAA34722	Aaa34722 Human ade
C 23	81.8	6.1	7033	3	AAF20844	Aaf20844 Human low
C 24	81.8	6.1	7033	10	ABZ96538	Abz96538 Human nuc
C 25	81.8	6.1	7036	2	AAAX55274	Aax55274 Human fac
C 26	81.8	6.1	209273	3	AAF21437	Aaf21437 Human fac
C 27	81.8	6.1	209274	10	ABZ97131	Abz97131 Human enz
C 28	81.4	6.1	6668	6	ABL33697	AbL33697 Human imm
C 29	81	6.0	1839	4	AAH17145	Aah17145 Human CDN
C 30	80.6	6.0	901	6	ABQ13742	Abq13742 Oligonuc1
C 31	80.6	6.0	901	6	ABQ13743	Abq13743 Oligonuc1
C 32	80.6	6.0	1343	6	ABZ78056	Abz78056 Human bre
C 33	80.6	6.0	3259	6	ABZ78057	Abz78057 Human bre
C 34	80.4	6.0	1236	3	AAA02163	Aaa02163 Human col
C 35	79.4	5.9	887	8	ABT42734	Abt42734 Human neu
C 36	79.4	5.9	887	8	ABQ20942	Abq20942 Oligonuc1
C 37	78.8	5.9	1995	6	ABQ20943	Abq20943 Oligonuc1
C 38	78.8	5.9	1995	6	AAI94064	Aai94064 Human neu
C 39	78.4	5.8	863	8	ABT42751	Abt42751 Human neu
C 40	78.4	5.8	863	8	ADL40874	Adl40874 Human ova
C 41	77.2	5.7	1041	5	ABQ75278	Abq75278 Human lun
C 42	77	5.7	876	6	ADC87306	Adc87306 Human GPC
C 43	77	5.7	1186	10	ADC87492	Adc87492 Human GPC
C 44	77	5.7	1225	10	ABV24723	Abv24723 Human pro
C 45	76.2	5.7	5242	5	ABV24723	Abv24723 Human pro

ALIGNMENTS

RESULT 1
AAF28296
ID AAF28296 standard; RNA; 2565 BP.
XX
AC AAF28296;
XX
DT 05-APR-2001 (first entry)
XX
DE LPL mRNA.
XX
KW Lipoprotein lipase; LPL; S447X; cardiovascular; diabetes; obesity;
KW deficiency; ss.
XX
OS Homo sapiens.
XX
PN WC200100220-A2.
XX
PD 04-JAN-2001.
XX
PF 23-JUN-2000; 2000WO-CA000762.
XX
PR 24-JUN-1999; 99EP-00202048.
XX
PA (UYBR-) UNIV BRITISH COLUMBIA.
PA (AMST-) AMSTERDAM MOLECULAR THERAPEUTICS BV.
PA (UYAM-) UNIV AMSTERDAM ACADEMIC HOSPITAL.
XX
PI Hayden MR, Kastelein JJP;
XX WPI; 2001-112388/12.
XX
PT Use of lipoprotein lipase S447X therapeutic for the preparation of a
PT pharmaceutical composition for the treatment of coronary heart disease,
PT atherosclerosis, angina pectoris, hypertension, diabetes, cachexia and
XX obesity.
XX
PS Claim 5; Fig 4; 5lpp; English.
XX
CC The present invention relates to use of a lipoprotein lipase (LPL) S447X
CC therapeutic for the preparation of a pharmaceutical for the treatment of
CC an LPL-responsive condition. Diseases that may be treated are complete
CC LPL deficiency, chylomicronemia, hyperlipidemia, partial LPL deficiency,
CC pancreatitis, hypertriglyceridemia, hypolipidemia, hypolipidemia (low high
CC density lipoprotein-cholesterol), cardiovascular disease, coronary heart

CC disease, coronary artery disease, atherosclerosis, angina pectoris,
CC hypertension, cerebrovascular disease, coronary restenosis, peripheral
CC vascular disease, diabetes, cachexia or obesity
XX
SQ Sequence 2565 BP; 1020 A; 739 C; 806 G; 0 T; 0 U; 0 Other;
Query Match 100.0%; Score 1344; DB 4; Length 2565;
Best Local Similarity 100.0%; Pred. No. 9.4e-278;
Matches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAGACACAGCGAGGACACCGCCACCCACCGGAGAGAGAGCGCGGACCCGACCAACACAG 60
Db 256 AAGACACAGCGAGGACACCGCCACCCACCGGAGAGAGAGCGCGGACCCGACCAACACAG 315
QY 61 CAGCAAAACCCAGGAGCAGCGGACCGGAACAGGAGAGAGAGGCGGCGCCAAACCGGCGG 120
Db 316 CAGCAAAACCCAGGAGCAGCGGACCGGAACAGGAGAGAGAGGCGGCGCCAAACCGGCGG 375
QY 121 CCCGACAGAGAGAACACAGACCCCAAGCAGGCGGACCGGCGGACCGGCGGACCGGCGG 180
Db 376 CCCGACAGAGAGAACACAGACCCCAAGCAGGCGGACCGGCGGACCGGCGGACCGGCGG 435
QY 181 GCCCGCGGACACCAACCGGCGGACAGAGGCGGCGGACCAACCGGAGGAGGAGACAC 240
Db 436 GCCCGCGGACACCAACCGGCGGACAGAGGCGGCGGACCAACCGGAGGAGGAGACAC 495
QY 241 CCCGACAGAGCAGCGGAGCAGCGGACCGGCGGACCGGCGGACCGGCGGACCGGCGG 300
Db 496 CCCGACAGAGCAGCGGAGCAGCGGACCGGCGGACCGGCGGACCGGCGGACCGGCGG 555
QY 301 AAGCAACAGAAACCGCGGACCGGCGGACCGGCGGACCGGCGGACCGGCGGACCGGCGG 360
Db 556 AAGCAACAGAAACCGCGGACCGGCGGACCGGCGGACCGGCGGACCGGCGGACCGGCGG 615
QY 361 AGCAGAGAGCGCACACACACAGAGGCGGCGGACCGGCGGACCGGCGGACCGGCGG 420
Db 616 AGCAGAGAGCGCACACACACAGAGGCGGCGGACCGGCGGACCGGCGGACCGGCGG 675
QY 421 GCAGCAGACACCGGAGGAGGACCGGCGGACCGGCGGACCGGCGGACCGGCGGACCGG 480
Db 676 GCAGCAGACACCGGAGGAGGACCGGCGGACCGGCGGACCGGCGGACCGGCGGACCGG 735
QY 481 AGGACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540
Db 736 AGGACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 795
QY 541 GAAACCAAGAGAGGCGCACAGGCGGACCGGCGGACCGGCGGACCGGCGGACCGGCGG 600
Db 796 GAAACCAAGAGAGGCGCACAGGCGGACCGGCGGACCGGCGGACCGGCGGACCGGCGG 855
QY 601 ACCGCGCAACACCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
Db 856 ACCGCGCAACACCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 915
QY 661 ACCGCGCAGGCGCACAAAGCCCAACCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 720
Db 916 ACCGCGCAGGCGCACAAAGCCCAACCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 975
QY 721 CAGGCGGAGACCGGAGCAGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780
Db 976 CAGGCGGAGACCGGAGCAGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1035
QY 781 AGACACACCCCAACACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 840
Db 1036 AGACACACCCCAACACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1095
QY 841 AGCGGCGAGAGGCGGAGCAGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 900
Db 1096 AGCGGCGAGAGGCGGAGCAGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1155
QY 901 AAAAGGAGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 960
Db 1156 AAAAGGAGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1215

QY 961 ACAGCCGAAAGACAGCGGCGGAAACACAGGCGGAAACACAGGCGGAAACACAGGCGG 1020
Db 1216 ACAGCCGAAAGACAGCGGCGGAAACACAGGCGGAAACACAGGCGGAAACACAGGCGG 1275
QY 1021 AGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
Db 1276 AGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1335
QY 1081 GCCCACCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140
Db 1336 GCCCACCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1395
QY 1141 CACACCCCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200
Db 1396 CACACCCCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1455
QY 1201 GAAAAAGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
Db 1456 GAAAAAGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1515
QY 1261 GAGCGCCCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320
Db 1516 GAGCGCCCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1575
QY 1321 AGCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1344
Db 1576 AGCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1599

RESULT 2
AAA10594
ID AAA10594 standard; DNA; 10732 BP.
XX
AC AAA10594;
XX
DT 29-JUN-2000 (first entry)
XX
DE Gene encoding a subunit of cellulose synthase.
XX
KW Cellulose synthase; cellulose production; increase yield; ds.
XX
OS Vigna angularis.
XX
PN JP2000060568-A.
XX
PD 29-FEB-2000.
XX
PF 26-AUG-1998; 98JP-00239998.
XX
PR 26-AUG-1998; 98JP-00239998.
XX
PA (MIZU/) MIZUNO K.
XX (OJIP) OJI PAPER CO.
XX
DR WPI; 2000-342371/30.
XX P-PSDB; AAY85179.
XX
PT A gene encoding a cellulose synthetic equipment - for the improvement in
PT the amount of cellulose synthesized in a plant body.
XX
PS Claim 2; Page 14-21; 32pp; Japanese.
XX
CC This sequence represents a gene encoding a subunit of the cellulose
CC synthase complex of Vigna angularis. The invention relates to subunits of
CC cellulose synthetic equipment, that can be used to increase the amount of
CC cellulose synthesized by a plant. The proteins and genes encoding them
CC can also be used to improve the properties of the cellulose being
CC produced by a plant
XX
SQ Sequence 10732 BP; 3149 A; 1212 C; 2074 G; 2046 T; 0 U; 2251 Other;
Query Match 7.7%; Score 103; DB 3; Length 10732;

Best Local Similarity 15.3%; Pred. No. 2.7e-12;
Matches 207; Conservative 522; Mismatches 611; Indels 12; Gaps 4;

QY	4	ACACAGCAGGACACGCCACCCACCGGAGAGAGAGCGCGGACCGCACAACACAGCAG 63		Db	10135	GNHASTRASGTHRAGSRASGYCGASNGTHRHSHVAGYSGARGHSRAFGGASNSRSGNV 10194
Db	9115	ASRYSGHYSGAGMTVAGGNSRGTHRSRYGARGAAYSAAGVAASASNSRHSNYSR 9174		QY	1073	AAAAACCCAGCCCGGAGGAGAGACAGCCAGCACAAGGCAACAAAGGGCAACAAAGGGAAG 1132
QY	64	CAAAACCCAGGACACGAGCGGACGGAACAGGAAGAGAGAGGCGGCCAAACCGCGCGCC 123		Db	10195	ATHRTHRSRCYSGYHSASNGSNVSGTYRTYRASNGYSGHTRHRSYASYSASYSRAAGSRGN 10254
Db	9175	HAAARGGASNGVARGAGGNGTHRNAAGAGSHSGYAAMTSRARGGSRMTGYASASNT 9234		QY	1133	GCCAAAAGCACACCCCAACGAAAGACAGGAGACAGAAAGCGCGCGGAGAAAGAAAG 1192
QY	124	GACAGAGAACACAGACCAAGCAGGGGACGCGCACCGGCCAGGAGCAACCCAGGCC 183		Db	10255	YSGVAGARGYSGNSRAATHRASARGGSRMT-ASNGNGNGSRGYSTRASSRARGGNSRH 10313
Db	9235	HRSRTHRTVYGAARGVAGYSGARGASNGGARGGHSRARGARGTHRAAGVARSYSYSGN 9294		QY	1193	AGCGCAGAGAAAAGGCCCCAGCGGAGGCGGCAAGCCAAAGGAGAAACCCCAAGGAAAG 1252
QY	184	GCGGCACACCAACCGGGGACAGAGGCGCGGACAGCAACGAGGAGGAGGAGAACACCC 243		Db	10314	YSGSRASNGYNGASNGYSASNGNYSASNSRASNGHSTRHGAAGYSAAHASNSY 10373
Db	9295	NGSRTHRYSYAAYSGCYSGNARGVAAARGYSARGRGYRAAAAAAYSMYASNGV 9354		QY	1253	GACCGGACGAGCGCCCAACACCAAGCCCAAGAAACAGAAAGCCCGCGGACAGCAACA 1312
QY	244	GGACAGCCACCGGGAACGCGGAGCCCGGAGCGGAGAGAGAGAGAGAGAGAGAGAG 9354		Db	10374	SVASRGGGASNYASNSRASNCYVGGAAYSCYSGNGNGSRMTSYSGCYSSRYSHA 10433
Db	9355	AGMTGYARGASRHGARGARGSRYSRTHRASNSRVAGSRVAASTHRSGTHRRTHR 9414		QY	1313	CCACACAGCAG 1344
QY	304	CAACAGAACCGCCGACCGACGAGACCAACGAGAGACAGAGAGAGAGAGAGAGAGAG 9474		Db	10434	SASYSASYSRARGASNASTRGTHRAAAASRGY 10465
Db	9415	YSASNTHRTHRGNCYSAAVAGGASNYSAVSGSRASNGYASNGNSRARGVAMT 9474				
QY	364	AGAGAGCGCACACACACAGAGGCGCGGACAGAGAGAGAGAGAGAGAGAGAGAGAG 9474		RESULT 3		
Db	9475	THRARGTHRAASRYSHSGSRNGASHSYSGVHSHSRVAGHSHSRASNSRASNSRHSRH 9534		RAV21511		
QY	422	--CAGGACAAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 9534		ID	AAV21511	standard; DNA; 6755 BP.
Db	9535	AASRSRAAAGYSRASASYSYTHRSYSAAGSRTAASRAASRASHSHARGSGYSY 9594		XX	AAV21511;	
QY	480	GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 539		XX	AC	
Db	9595	SGRSYCSYSSRVAGYAAASASASNTASASHVAGMTGYSAAVASRVAGYSAATHRG 9654		XX	AC	
QY	540	AGAAACCAAGAGGCGCACAGGCGAGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAG 599		DT	17-OCT-2003	(revised)
Db	9655	SRSRASRYSAVATHRGASNSRGTHTHYTHRSYSGSRTHRGGYSYSGRVARASHS 9714		DT	17-AUG-1998	(first entry)
QY	600	AACCGCGCAACACCGGACAGACAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 659		XX	Staphylococcal bacteriocin BacRI operon.	
Db	9715	RGNTHRTHRTTHRRGVAASMTGRTMSRASHSBSASRYSSRYASNTHRCSYSSRA 9774		DE	BacRI; bacteriocin; antimicrobial; antibacterial; antibiotic;	
QY	660	GACCGCAGACCCACAAAGCCCAACAGAAACACAGGAGGAGGAGGAGGAGGAGGAG 719		KW	Moraxella bovis; infectious bovine keratoconjunctivitis; cancer; therapy;	
Db	9775	SHYSGYTHRVARGYTRGNASVAVAYSTRHGNASHSVATHRHSYSGYRASASHSASYSV 9834		KW	ds.	
QY	720	ACAGCGCGAGACCGGAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 779		XX	Staphylococcus aureus; strain UT0007 (ATCC 55800).	
Db	9835	AAAARGHVAASNAASHASHASRSRARGGYRVASHASTHRGNASNRNGCYSHSCY 9894		XX	WO9812319-A1.	
QY	780	AGACACACCCCAACACAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 839		PD	26-MAR-1998.	
Db	9895	SSRCYSASRSRVAASASNSRAAGYASNAASRYASRGSRGNGASRYSSRY 9954		XX	18-SEP-1997; 97WO-US016758.	
QY	840	CAGC----GCGAGCGGAGGAGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 895		PF	19-SEP-1996; 96US-00710561.	
Db	9955	YSGVAGARGAAVAASTYRASRSRASRNGYSGYVARGASHSTRYVSASNGYMTTRTH 10014		PR	17-SEP-1997; 97US-00931999.	
QY	896	ACCAGAAAAGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 952		PA	(UNIV) UNIV KANSAS STATE RES FOUND.	
Db	10015	GYTMTVAARGVAGHNTYSTRSGRSYSTRVAGGHSVACYSTVTRASSRGYSTRAST 10074		XX	Iandolo JJ, Crupper SS;	
QY	953	AAAGCAGACAAAGCCGAAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1012		XX	WPI; 1998-230316/20.	
Db	10075	YRGVASNHAAGVTHRTTHRAAASSTRMTASHNSHCYSRSGNVAASRMTARGASAYS 10134		XX	Therapeutic proteinaceous substances from Staphylococcus aureus - useful	
QY	1013	GGAACGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1072		PT	to inhibit growth of wide range of prokaryotic or eukaryotic cells, e.g.	
				PT	Moraxella bovis causing infectious bovine keratoconjunctivitis.	
				PS	Claim 2; Page 19-23; 38pp; English.	
				CC	This polynucleotide comprises the bacteriocin BacRI operon of	
				CC	Staphylococcus aureus UT0007. The sequence of the BacRI operon was	
				CC	determined by N-terminal sequencing of purified BacRI peptide (see	
				CC	AAW54171), with back-translation and plasmid analysis. The BacRI operon	
				CC	includes the BacRI gene (see AAV21510), a homologue of the cym gene of	
				CC	the cytolytic operon of Enterococcus faecalis whose function is involved	
				CC	in the maturation of pre-cytolysin, an ATP-transporter gene, bioI and	
				CC	bio2 genes related to lactococin biosynthesis and modification, and a	
				CC	gene involved in immunity function. BacRI peptides can be produced by	
				CC	construction of an expression vector containing an oligonucleotide or	
				CC	operon coding for BacRI, and use of the vector to transform host cells	
				CC	for BacRI expression. The entire BacRI operon has been cloned into	

CC plasmid pUB110, and Bacillus subtilis transformants secreted the
 CC recombinant BacRi peptide into the medium. Bacteriocin BacRi is active
 CC against many Gram-positive and Gram-negative organisms such as Bordetella
 CC bronchoseptica, Pasteurella multocida and Staphylococcus aureus;
 CC Moraxella bovis, causing infectious bovine keratoconjunctivitis, is
 CC especially sensitive. BacRi can also be used as an anti-cancer agent.
 CC (Updated on 17-OCT-2003 to standardise OS field)
 XX
 SQ Sequence 6755 BP; 3903 A; 1252 C; 1600 G; 0 T; 0 U; 0 Other;
 Query Match 7.6%; Score 102; DB 2; Length 6755;
 Best Local Similarity 45.3%; Pred. No. 4e-12;
 Matches 568; Conservative 0; Mismatches 675; Indels 11; Gaps 5;
 52 ACACACAGCAGCAGCAAAACCCAGGACCCAGGCGGACGGAACAGGAGAGAGAGGGGCGCAA 111
 5413 ACAAGGGGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 5472
 112 AACGGGCGCGCGACAGAGAGAACACAGACCCAGAGAGGGGGAGCGCGACGGCCAGGA 171
 5473 AAAGGGCAAGCGAAGAGAGGAGAAACAGAAACCAAGCGACAGAGAGAAAGAAAGCGGAAC 5532
 172 GCAACCCAGCGCGCGGACACCAACCGGGGACAGGAGGCGCGGACACAGAGGAGGA 231
 5533 CCACAGCGGGAGGGGAAACAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 5592
 232 GGAGAACACCCCGGACAGGACCCAGCGGAACAGCCGCGAGCCCGCGGACGAGCAAGCGG 291
 5593 GGGGCAACACAGACAGCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 5652
 292 ACCAAAGAAAGCAACAGCAACCGCGGACCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 351
 5653 ACACAAAGAAAGCAACAGCAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5710
 352 CGCCCGGAGACAGAGACGACACACACACAGCCGAGGAGGAGGAGGAGGAGGAGGAGGAG 411
 5711 ACAACCAAGCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 5770
 412 AAACCGAGGAGGAGCAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 471
 5771 CGAGGACCGGACAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5830
 472 AGCAG 531
 5831 ACACCCAGAAACACCAACACAGCAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5890
 532 CGGGAAGAGAAACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 591
 5891 AGGGAAGAGGAGAAACCAAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5950
 592 GGAGAAAGAAACCGCGCAACACCGGCGAGAGACAAAGAAAGAGAGAGAGAGAGAGAGAGAG 651
 5951 AGAAGAGGAAAGAAACACGAAACGGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAGAG 6010
 652 AGACCGAAGAGCGCGAGAGCCCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 710
 6011 ACGAGAAAGAAAG 6070
 711 --ACCCACCAACAGCGCGAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 768
 6071 GGAGAGAAAG 6130
 769 GAAGCCACAAAG 828
 6131 GAAGGAG 6190
 829 GAAG 888
 6191 ACCAAAGAAAG 6249
 889 AGGAG 946
 6250 AAAAAACGAG 6309

QY 947 CGGAGGAGAGCCAGACAGCCGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1006
 DB 6310 CACAAAG 6369
 QY 1007 ACGGAG 1066
 DB 6370 AACAGCAAAAG 6429
 QY 1067 GACCGAAAG 1126
 DB 6430 AAAAAAG 6489
 QY 1127 GGAAGGAG 1186
 DB 6490 ACAGCAAAAG 6546
 QY 1187 AAAAG 1246
 DB 6547 GAG 6606
 QY 1247 GGAAGGAG 1300
 DB 6607 AACACGAG 6660
 RESULT 4
 ABX11086/c
 ID ABX11086 standard; DNA; 42999 BP.
 XX
 AC ABX11086;
 XX
 DT 24-APR-2003 (first entry)
 XX
 DE Human ribosomal RNA (rRNA) gene.
 XX
 KW Nucleic acid amplification; artificial chromosome isolation; MAC;
 KW DNA delivery; mammalian artificial chromosome; gene therapy; organ;
 KW humanised genetically transformed animal; chromosomal element;
 KW gene product production system; transgenic; centromere function;
 KW information storage vehicle; artificial chromosome vector; human;
 KW species-specific artificial chromosome; ribosomal RNA; rRNA; gene; ds.
 OS Homo sapiens.
 XX
 XX US2002160410-A1.
 XX
 XX 31-OCT-2002.
 XX
 XX 17-APR-2002; 2002US-00125767.
 XX
 XX 10-APR-1996; 96US-00629822.
 XX 15-JUL-1996; 96US-00682080.
 PR 07-AUG-1996; 96US-00695191.
 PR 10-APR-1997; 97US-00835682.
 PR 28-NOV-2000; 2000US-00724693.
 XX (HADI/) HADIACZY G.
 PA (SZAL/) SZALAY A. A.
 XX
 XX Hadlaczky G, Szalay AA;
 XX WPI; 2003-265757/26.
 DR
 XX Amplifying nucleic acid for constructing artificial chromosomes,
 PT comprises introducing a DNA fragment into a cell to incorporate the DNA
 PT into a chromosome, and identifying cells with chromosomes that has
 PT undergone amplification.
 XX
 XX Disclosure; Page 76-94; 107pp; English.
 PS
 XX The present invention relates to methods for amplifying nucleic acids in
 CC cells, for isolating artificial chromosomes, and preparing cell lines

QY	96	AGAGAGAGCGGGCCCAAAACGGGGCCGCCGACAAAGAGAGAACACAGACCCCAAGCAGGGGGAC	155
DB	25972	AGAGAGAGAGAGAGAGACAGAGAGAGACGAGAAAAAAGAGAGAGGCGAGACNNNNAGAG	25913
QY	155	GGCGCAGCGGGCCAGGAGCAACCCAGCGCGCGGCGCACACCAACCGGGGACAGAGGGGCC	215
DB	25912	AAAGACAGAGACAGACAGAGAGAAACAGCGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG	25853
QY	216	GGACAAACGAGAGAGAGGAGAGAACACCCCGGACCAAGCCACCGCGAAACACCCGAGCCCGAGC	275
DB	25852	AGTAAAGAGAGAGAGAGAGAGACAGAGAGAGAGAGAGAAATAGACAGAAAGANNNNACAA	25793
QY	276	GGCGCAGCAGGAGCGACCAAAAGAAAGCAACAGAAACGGGCCGACACAGCGGACCAACGA	335
DB	25792	GTGAGAG-AGGCAGAGAGAGAGAGAGAGAGAGAGAGAGACAGAGACAGAGAGANNNGAG	25734
QY	336	GAGCAGAAAGCCCGAGCGCCCGAGAGCAGAGAGCGCACACACACACAGAGAGGGCCCCG	395
DB	25733	AAAGAGACAGACAGACAGAGAAAGAGACAGACAGAGAAAGACAGAGAGAGAAANNNGAA	25674
QY	396	CGAAGCAGGAAACACAGAAACACGAGGCGCAGGACAAACCCGAGAGAGACACAGCCAGGAGAC	455
DB	25673	ACAGGCAGAGAGAGAGAGAGCTAGCGAGAGAGAAACAGAAAGGTAGGAGAGANNNGAG	25614
QY	456	AGGAGAGACACCGGAGCAGAGAGAGAGAGCGAGAGGAGCCAGCAGGAAAGCCCCACGAG	515
DB	25613	AGACAGACAGACAGATGGACAGCGCAGAGAGAGAGAGTAAAGACAGACAGACAGACNNNTGA	25554
QY	516	CGCCACACCCACGACCCGGAAGAAGAAAAACCAAGAAAGGCCACAGGGCCACCCAGGAGAC	575
DB	25553	GAGAGACAAAGAGAGAGAGAGAGAGAGACAGAGACAGAGACAGAGACAGACAGACNNNGAA	25494
QY	576	CGAGAAAGGGCCGAGAGAGAAAGAACCGGCGCAACACGGGCGAGAGACAAAAAGCAGAGC	635
DB	25493	AGAAAGAGAGAGACAGACAGACAGAAAAAGACAGAGACGCGACAGAGAGAGACANNNTAAC	25434
QY	636	CAAAAGAGCAGCAAAAGACCGAAGACCGCCAGGCCCAAAAGCCCAACCAAGAAAGAC	695
DB	25433	AGAGAGAAAGAGAAAGACAGAGAGAGCGGAGAGAGAGAGAGAGAAACANNNGGC	25374
QY	696	ACGGGACCGAGAGGAAAAACCAACCAACAGGCCGAGACCGAGGCCACCGGCCGAGAGAGAA	755
DB	25373	AGGGAGAGAGACAGAGAGAGACAGACAGATAGACAGCAGAGAAAGAGAGTAAANNNGAA	25314

25442	CAGAGAAACAGAGAGAAGAGAGAAACACAGAGAGCGGAGAGAGAGAGAGAGAGAA	25387
587	AAGAAAGACACGGGACGAGAGGAAAACCCAAACCACAGGCCGAGCACCGAGGCACCGGCCG	746
25382	CAGAAAGCGCAGGAGAGAGACACAGAGAGAGACACAGATAGACAGGCAGAGAAAGAGGT	25323
747	AGAGGAGAACACCCACACACCGCCGAAGCCACAAAGAACCCACCCCACACAGAGAGAA	806
25322	AAGACAGAAAGTAGGCAACAGAGAGAGACAGAGAGACAGAGAAAGAGAAAGAGAGAGGC	25263
807	GGAGAACACACAGAGAGCCAAAGAGAGAGACACAGCGGCAGACCGGGACAGCCCCGCC	866
25262	AGACACAGACAGAAAGGGACAGACAGAGAAAGACAGAGACAGAGAGAGAGAG-	25213
867	GCCACAGAGACAGAGAAAAAGCAGGAGAGACCCAGAAAAAGGGACCCGACGGGAGAGAGGCC	926
25212	-----AGAGAGAGAGAGAGAGAGAAACACAGACAGAAAGAGAGAGAGAGAG	25159
927	AGCAGAAAGGAAAGGCCACCGCGGAGGAAAGCCAGACAAGCCGAAAAGAGAGAGAGAG	986
25158	AGACAGACAGACAGACCGACAGGCAGAGAGAGAGAGTAAAGACAAAAGATTACACACAGA	25099
987	GGCGAACACACAGAACAAAGAACGCCAGGNACGGAAAGAGAGAGGAGGAGAGAACACAAAC	1046
25098	GCGACAGACACAGAGAGAGAGAGACAGAGACAGAGACAGAGAGAGAGATAAAGACAGACAC	25039
1047	AACCCAGGGGGCAAAGAGGACCCGAAAAACCCAGCCCCACCCGAGAGAGGAGACAGCCAAG	1106
25038	GAGAGAGACATGTAGAAACACAGACAGAAAGAGAGGGGACAGAGAGAGAACACACAGAGAG	24979
1107	CACAAAAAGGCGCAACAAAGGGGAAAGGGCCAAAAGCACACCCCAACGAAACACAGGGACA	1166
24978	GNNNAGAG- - - - -AGACAGAGAGAGAGAGAAAGCAGAAAGGGAGGAGACACAGAGAGA	24926
1167	GAAGAGGGCGCCAGAAAGAAAGAGCCGACAGHAAAAAGGCC---CAGGCGAGGG	1223
24925	GACACACAGACAGGCAGAGAGAGAGAGTTAAGACAGAGATAGGCACAGAGAGAGAGAG	24866
1224	CCAAAGCCAAAGGAGAACCCCAAGGAACGGACCGGACGAGGCC---CAAAACACCAAGCC	1280
24865	ACAGACAGAGAGACACAGAAAGAAAGAGAGGGGCACACAGACAGAGAAAGAGACAGAT	24806
1281	CAAGAAACAGAAAAAGCCCGCGCACGAAACAGACCCACACACAGAGAGAGAGAGA	1335
24805	AGAGAAAGAGAGAGGCGACAGACAGAGAGAGACAGAAAACAGACAGAAAGAGAGAGA	24751
RESULT 8		
ABS65032/c	ABS65032 standard; DNA; 42999 BP.	
ID		
XX		
AC		
XX		
DT	(first entry)	
XX		
DE	Invertebrate foraging behaviour associated human DNA sequence #12.	
XX		
KW	Attention deficit hyperactivity disorder; ADHD; hypertension;	
KX	invertebrate foraging behaviour; nitric oxide; hypotensive;	
KW	cGMP-dependent protein kinase; human; neuroleptic; ds; gene therapy.	
XX		
OS	Homo sapiens.	
XX		
FN	WO200259370-A2.	
XX		
PD	01-AUG-2002.	
XX		
PF	13-DEC-2001; 2001WO-US048087.	
XX		
PR	15-DEC-2000; 2000US-00738630.	
XX		
PA	(NEUR-) NEUROSCIENCES RES FOUND INC.	
XX		

[illegible]

XX DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 25802.

XX KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;

XX KW drug; side effect; cancer; central nervous system; cardiovascular;

XX KW gastrointestinal; respiratory system; single nucleotide polymorphism;

XX KW SNP; cell differentiation; ds.

XX OS Homo sapiens.

XX KW WO200218632-A2.

XX PN 07-MAR-2002.

XX PD 01-SEP-2001; 2001WO-EP010074.

XX PF 01-SEP-2000; 2000DE-01043826.

XX PR 05-SEP-2000; 2000DE-0104543.

XX PA (EPIG-) EPIGENOMICS AG.

XX PI Olek A, Piepenbrock C, Berlin K, Guetig D;

XX DR WPI; 2002-371829/40.

XX KW Determining the degree of cytosine methylation in genomic DNA, useful for

XX PT diagnosis and prognosis, comprises selective hybridization of amplicons

XX PT from chemically treated DNA.

XX PS Claim 12; 56pp + Sequence Listing; 56pp; German.

XX CC This invention describes a novel method for determining the degree of

CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a

CC genomic sample of DNA. The sample is treated chemically to convert

CC cytosine (C) but not methylated C to uracil, then part of the genomic

CC DNA that contains the target C is amplified to form a labeled amplicon.

CC The amplicon is hybridised to two classes, each with at least one member,

CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the

CC degree of hybridisation to both classes is determined from the label on

CC the amplicon. From the ratio of labels hybridised to the two classes of

CC oligomers, the degree of methylation is calculated. The method is used:

CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs

CC and of a wide range of diseases, e.g. cancer, disorders of the central

CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,

CC particularly by detecting mutations or single nucleotide polymorphisms

CC (SNP's); and (ii) for differentiation of cell or tissue types and for

CC investigating cell differentiation. The method allows the methylation

CC status of many C residues to be determined simultaneously. ABQ13410-

CC ABQ34121 represent genomic DNA sequences used to illustrate the method

CC for determining the degree of cytosine methylation described in the

CC disclosure of the invention

XX SQ Sequence 1200 BP; 919 A; 153 C; 112 G; 0 T; 0 U; 16 Other;

Query Match 6.6%; Score 88.6; DB 6; Length 1200;

Best Local Similarity 43.8%; Pred. No. 2e-09;

Matches 349; Conservative 0; Mismatches 447; Indels 0; Gaps 0;

QY 536 AAGAGAAACCAAGAGGCCACAGGGCAGGCAAGCAAGAGCCGAGAAAGGCGCGAGGAG 595

DB 192 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACGACGCGCGGCAAAACGCCCAANAC 251

QY 596 AAGAACCGCGCAACACGGCGAGAGACAAAAAGCAGAGCCAAAGAGCAGCAAGAGAC 655

DB 252 GAAAAAAAAACGCAAAAAAAAAACAAANAAAAAAAAAAAAAAAAACGAAACCGCGCAANA 311

QY 656 CGAGAGCCGCGAGGCCCAAGCCCAACCAAGAAAGACACGGGACGAGAGAAACCCA 715

DB 312 CGAAACGAACGMAAAAAAAAAACGCAACAAACAAACAAACGACGCGACAAACGAANAA 371

QY 716 ACCAACAGGCGGAGCCGAGGACCGCGCGGAGGAGAGACACCCACACCGCGGAGGCCA 775

DB 372 AAAAAAAAAAAAAAAAAACGCGCAACAAACGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 431

QY 776 CAAAAACACCCCGAAACACAGAGGAGAGAACACACGAGAGCCCAAGGAGAGG 835

DB 432 AAAAAAAAAACGAAAAAACAAGAAAAACGAAAAAACAAGAAAAAACAAGAAAAA 491

QY 836 ACAACACGCGGAGAGCGGGAGAGCCCGCGCCACAGAGAGACAGAGAAAGCAGAGAG 895

DB 492 AAAACGAAACGCGCAAAAAAAAAACGAAAAAACAAGAAAAAACAAGAAAAAACA 551

QY 896 ACCAGAAAAAGGACCGCAGGGAGAAAGCGCAGCAGAGAAAGGACCGCGAGGAAA 955

DB 552 GAAAAAAAAAAAAAAAAACGAAAAAACAAGAAAAAACAAGAAAAAACAAGAAAAA 611

QY 956 GCCAGACAAGCGGAAAGAGAGCGGCGAAACCGGCGGAAACACAGAAACAAAGACGCGAGGA 1015

DB 612 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACGAAACGCGMAAAAAAAAAACA 671

QY 1016 ACGGAAGAGAGAGGAGAGAAACACAAAAACAACCCAGGGGGGGGCAAAAGGAGCGGAAAA 1075

DB 672 AAAAAAAAAACGAAACGAAAAACGCGAAAAACACAAACGAAAAAACAAGAAAAA 731

QY 1076 ACCCAGCCCGCAGAGAGGAGCAGCCAGCACAAGAAAGGCGCAACAAAGGGAGGGGCC 1135

DB 732 AAAAAAAAAACCGAAACCGACGCGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 791

QY 1136 AAAAGCACACCCCAACGAAAAAGACAGGACAGGACAGAAAAAGCGCGCCGAGAAAAAGAGGAGC 1195

DB 792 AAAAAAAAAACGAAACCGAAACCGCAGACCGCGGACGAAAAAACAAGAAAAAACAAG 851

QY 1196 GCAGAGAAAAAGGCCCGCGAGGGCCAAAGCCAAAGGAGAGACCCCAAGAGACGCGC 1255

DB 852 AAAAAAAAAAAAAAAAAACGAAAAAACAAGCGACGCGAAAAACGAAAAACGAAAAACGC 911

QY 1256 CGGACGAGGCGCCCAACACACCAAGCCCAAGAAACAGAAAGCCCGCGGACGACAGACCA 1315

DB 912 GAAAAAAAAACAAACGAAACCGAACCGAACCGAAAAAACAAGAAAAAACAAGAAAAA 971

QY 1316 CACACAGCAGAGAGAAGA 1331

DB 972 AAAAAAAAAAAAAA 987

Search completed: October 21, 2004, 04:14:57

Job time : 739 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 21, 2004, 03:30:20 ; Search time 137 Seconds
(without alignments)

6972.995 Million cell updates/sec

Title: US-10-019-341-4_COPY_256_1599

Perfect score: 1344

Sequence: 1 aagacacagcgaggacagc.....gagaagagagaagacccgc 1344

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents NA.*

1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*

2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*

3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*

4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*

5: /cgn2_6/ptodata/1/ina/PCFUS_COMB.seq.*

6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	129.6	9.6	7218	1	US-08-232-463-14
C 2	102	7.6	6755	3	US-08-931-999-4
C 3	82.8	6.2	851	4	US-09-270-767-8286
C 4	82.8	6.2	851	4	US-09-270-767-23568
C 5	82.6	6.1	1827	4	US-09-270-767-1308
C 6	82.6	6.1	1827	4	US-09-270-767-16590
C 7	70.8	5.3	1298	3	US-08-948-705-3
C 8	70.8	5.3	1298	4	US-09-510-543-3
C 9	67.2	5.0	90541	4	US-09-759-359A-3
C 10	67.2	5.0	90541	4	US-10-207-973-3
C 11	63.8	4.7	53526	3	US-08-658-136-2
C 12	63.8	4.7	53577	3	US-08-658-136-1
C 13	63.6	4.7	705	4	US-09-270-767-5061
C 14	63.6	4.7	705	4	US-09-270-767-20343
C 15	63.6	4.7	1835	4	US-09-216-393B-80
C 16	63.4	4.7	1923	4	US-09-134-000C-466
C 17	62.4	4.6	1374	4	US-09-270-767-2575
C 18	62.4	4.6	1374	4	US-09-270-767-17857
C 19	61.2	4.6	176373	3	US-09-128-155-17
C 20	60	4.5	152331	3	US-09-128-155-16
C 21	59.6	4.4	1138	4	US-09-509-712B-107
C 22	59.4	4.4	289	3	US-09-007-005-17
C 23	59.4	4.4	289	3	US-09-244-796-17
C 24	56.6	4.2	152331	3	US-09-128-155-16
C 25	56	4.2	72604	3	US-09-268-992-7
C 26	56	4.2	72604	3	US-09-657-474-7
C 27	55.8	4.2	319608	4	US-09-539-333D-1

28	55.8	4.2	319608	4	US-09-679-409-1	Sequence 1, Appli
C 29	55.6	4.1	936	4	US-09-270-767-4464	Sequence 4464, Ap
C 30	55.6	4.1	936	4	US-09-270-767-19746	Sequence 19746, A
C 31	55.6	4.1	1559	3	US-09-019-095A-7	Sequence 7, Appli
C 32	54.2	4.0	510	4	US-09-248-796A-13735	Sequence 13735, A
C 33	54.2	4.0	168174	4	US-10-071-411A-63	Sequence 63, Appli
C 34	54.2	4.0	168174	4	US-10-071-411A-2	Sequence 2, Appli
C 35	54	4.0	435	3	US-09-328-352-583	Sequence 583, App
C 36	53.8	4.0	4453	3	US-09-146-053-5	Sequence 5, Appli
C 37	53.6	4.0	774	4	US-09-252-991A-3254	Sequence 3254, Ap
C 38	53.4	4.0	202001	4	US-09-734-674-3	Sequence 3, Appli
C 39	53	3.9	432	4	US-09-328-352-585	Sequence 585, App
C 40	53	3.9	447	4	US-09-328-352-584	Sequence 584, App
C 41	52.6	3.9	668	4	US-08-516-545-1	Sequence 1, Appli
C 42	52.6	3.9	668	4	US-09-050-344-1	Sequence 1, Appli
C 43	51.8	3.9	6158	4	US-09-799-451-897	Sequence 897, App
C 44	51.6	3.8	336	4	US-09-248-796A-13866	Sequence 13866, A
C 45	51.4	3.8	893	4	US-09-509-712B-46	Sequence 46, Appli

ALIGNMENTS

RESULT 1
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-9300
; TELEFAX: (703) 683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-F1s
US-08-232-463-14

Query Match 9.6%; Score 129.6; DB 1; Length 7218;

RESULT 2
 US-08-931-999-4
 ; Sequence 4, Application US/08931999
 ; Patent No. 6043219
 ; GENERAL INFORMATION:
 ; APPLICANT: Iandolo, John J.
 ; APPLICANT: Crupper, Scott S.
 ; TITLE OF INVENTION: Broad Spectrum Chemotherapeutic Peptide
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Hovey, Williams, Timmons & Collins
 ; STREET: 2405 Grand Boulevard, Suite 400
 ; CITY: Kansas City
 ; STATE: Missouri
 ; COUNTRY: U.S.A.
 ; ZIP: 64108
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/931,999
 ; FILING DATE:
 ; CLASSIFICATION: 514
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/710,561
 ; FILING DATE: 19-SEP-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Collins, John M.
 ; REGISTRATION NUMBER: 26,262
 ; REFERENCE/DOCKET NUMBER: 25043-A
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 816/474-9050
 ; TELEFAX: 816/474-9057
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 6755 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: unknown

/	MOLECULE TYPE:	DNA (genomic)
/	HYPOTHETICAL:	NO
/	ANTI-SENSE:	NO
/	ORIGINAL SOURCE:	
/	ORGANISM:	Staphylococcus aureus
/	STRAIN:	UT0007
/	US-08-931-999-4	

		Query Match	7.6%;	Score 102;	DB 3;	Length 6755;
		Best Local Similarity	45.3%;	Pred. No. 9.1e-14;		
Matches	568;	Conservative	0;	Mismatches 675;	Indels 11;	Gaps 5;

QY	52	ACAACACAGCAGCCAAACCAGGACACCGCGGACGGGAACAGGAAGAGAGAGAGGGGGGCCAA	111
DB	5413	ACNAGCGMAAATAAAAAAAGAATAAAAAAATAAAGAAATCAATACCGCGGGAAGCCAGGAAG	5472
QY	112	AACGGGCGCCCGCACAGAGAGAACCGAGACCAAGCAGGAGGGGACGGGCACCGGCCACGA	171
DB	5473	AAAGGCAAGCGAAGAGGGGAAAACAGACTCAACGCGACAGAGAGAAAGGAACCGGAAC	5532
QY	172	GCATCCCAGGCCCGGGCACACCAACCGGGGACAGAGGAGGCCCGGCACACCGAGAGGAGA	231
DB	5533	CCCAAGCAGGAGGGGAAAAACAAAAAATAAAAAAATCGCTAAAAAATCAAAGAAGCGGA	5592
QY	232	GGAGAAACACCCCGGCAAGCCACCGGGACACAGCCGAGCCCCAGCCGGCACGAGAACGC	291
DB	5593	GGGGCACAAACAGACAGCAAGAAAGAAAAAATAAAAAATACAGGGGAACACAGGAGCGCGAACG	5652
QY	292	ACCAAAAGAAAGTAACAGAAACCGGCCACAGCGGACCAACGAGAGCAGAGAGCCCGAG	351
DB	5653	ACAAAAAATAAACAGCACAAA--AAACCAAGAAACACGACAAAGAGAGAAAAAGAAAAAG	5710
QY	352	CGCCCGCAGAGCAGAGACGACACACACACAGAGGGGCCCGCGGAAGCAGGAACCGAC	411
DB	5711	ACAAACCAAGCAATAGAAAAAGAAATACAAAGACAAACAGAAAGAAAAAGAAAAAATA	5770
QY	412	AAACACAGGGCGAGACAACCCGAAGAGGACACCGCAGGAGACAGAGACAGAGAGACCGCGG	471
DB	5771	CCAGCACCCCGCACAAAAAGAGAGCAGAGGGGAAACCGGACAAACACAGGAGCAGGGACA	5830
QY	472	AGCAGAGAGGAGCGGAGAGGACACAGCAGAGGAAGGCCCCACAGCGCGCCACACCCACGAC	531
DB	5831	ACACCCAGAAAACACCAACACACGACACCAAGAGAAAGCAAGAGAGCAAAAAACACAAAC	5890
QY	532	CCGGAAGAAATAACCAAGAAGGCCACAGGGCAGCGCAAGGAAGCCGAGAAAGGGCCGGA	591
DB	5891	AGGGAGAGGAAAAACCAAGAGCGAAGACAAAAAAGAAAGAGGCGACAGCAGAGAAAAACA	5950
QY	592	GGAGAAAGAACCCGCAACACCGGCAGAGACAAAGCAGAGCCAAAGAGAGCAGCAAA	651
DB	5951	AGAAAGGAAAAAATACCGACCGGAAAAAACAAGAGGAAATAGAGAACAAAAAGGAGA	6010
QY	652	AGACCGAAGACCCGACAGCCCCAACAAAGGCCAACCAAGAAAAAGACCGGACGAGGAAG-	710
DB	6011	ACNGHAAATACAAAGAAACAGGAACGAGCAAGAAAGAAACAAACAAAAAAGAGAGGAG	6070
QY	711	--ACCCAACCAACAGGCCGAGACCGAGGACACCGGCCGAGAGAGGAGAACCCACACCGCC	768
DB	6071	GGAGGGAAGAAAGGAGCAGGAAAAAGACCAAGAAAAACCAACCGAAAAAGACGCGAGA	6130
QY	769	GAAGCCACAAAAAGACCAACCCCAACACAGAGGAGAGGAGAAACACCAAGGAAGCCAAG	828
DB	6131	GAAGGAACGAAGAGGAAAGAGAGACCCAAAAAATAACAAACCGGACAGAGACAAAG	6190
QY	829	GAAGAGACAAACAGCCGACAGCGGGAGACGCCCCGCCCAACAGAGACAGAGAAAGC	888
DB	6191	ACCAAAAACCAACCGAAAAACGACAGAAACAAAAACAATAAG--AGAAAAAGAAAAAGA	6249
QY	889	AGGAGACACAGAAAAAGGACCGCAGGAGGAAGGCCAGCAGCAAAAGGAAG--CACCG	946
DB	6250	AAAAAAACGCAAGGAAACAAAGAGAGGGAAACAGCAAAAAAATCCGCAAAAGCAGCC	6309
QY	947	CGGAGGAAGCCAGACAAAGCCGAAAGAGACAGCGCGGAAACGGCGGAACACAGAACAAAGA	1006

Db 467 ATGAAGAGGGGAGAAAGAAAGGAAGAGAAAGAAAGAGCGGAAAGAAAGAGAAAGAG 408
QY 948 GGAGGAAGCCAGACAGCCGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1007
Db 407 GAAGGAAG 348
QY 1008 CGGAGGAAGCGAG 1067
Db 347 AAAAGGGGGAG 288
QY 1068 ACCGAG 1127
Db 287 AG 228
QY 1128 GAAGGGGCAAAAGCACACCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1187
Db 227 AAAGGAG 168
QY 1188 AAAGGAG 1209
Db 167 TGGGGGGGGGAG 146

RESULT 5

US-09-270-767-1308
; Sequence 1308, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270.767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1308
; LENGTH: 1827
; TYPE: DNA
; ORGANISM: *Drosophila melanogaster*
US-09-270-767-1308

Query Match 6.1%; Score 82.6; DB 4; Length 1827;
Best Local Similarity 48.4%; Pred. No. 1.4e-09;
Matches 320; Conservative 0; Mismatches 334; Indels 7; Gaps 3;
QY 534 GGAAG 593
Db 1158 GGAAG 1217
QY 594 AGAAG 653
Db 1218 AAG 1277
QY 654 ACCGAG 713
Db 1278 AAG 1277
QY 654 ACCGAG 713
Db 1278 AAG 1337
QY 714 CAACCAAG 770
Db 1338 GAAAG 1397
QY 771 AGCCCAAG 830
Db 1398 GAAAG 1457
QY 831 AGAGGAG 890
Db 1458 AAG 1517
QY 891 GAG 950
Db 1518 AAG 1574

QY 951 GGAAGCCAG 1010
Db 1575 AGAGGGAG 1633
QY 1011 CAGGAAG 1070
Db 1634 AAG 1693
QY 1071 GGAAG 1130
Db 1694 AGGAAG 1753
QY 1131 GGGGCAAG 1190
Db 1754 GAAG 1813
QY 1191 G 1191
Db 1814 G 1814

RESULT 6

US-09-270-767-16590
; Sequence 16590, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270.767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16590
; LENGTH: 1827
; TYPE: DNA
; ORGANISM: *Drosophila melanogaster*
US-09-270-767-16590

Query Match 6.1%; Score 82.6; DB 4; Length 1827;
Best Local Similarity 48.4%; Pred. No. 1.4e-09;
Matches 320; Conservative 0; Mismatches 334; Indels 7; Gaps 3;
QY 534 GGAAG 593
Db 1158 GGAAG 1217
QY 594 AGAAG 653
Db 1218 AAG 1277
QY 654 ACCGAG 713
Db 1278 AAG 1337
QY 714 CAACCAAG 770
Db 1338 GAAAG 1397
QY 771 AGCCCAAG 830
Db 1398 GAAAG 1457
QY 831 AGAGGAG 890
Db 1458 AAG 1517
QY 891 GAG 950
Db 1518 AAG 1574
QY 951 GGAAGCCAG 1010

Db	1150	GAAGAGAGAGAACCGCAGAGAGTGGAG-GAAGAGAGGCGGAGACAGAGAGAGGAAGA	1208
Qy	780	AAGACCACCCCAACACAGAGAGAGAGGAGAACACACAGGAAGCCAAAGGAAGAGGACAA	839
Db	1209	AGAGAGAGAACCGCAGGAGAGTGGAGGAGAGAGAGGCGGAGACAGAGAGGAGGAGAGAGA	1268
Qy	840	CAGCGGAGAGCGGAGAGAGCCCGCGGCCACAGAGACAGAGAAAAGCAGGAGAGACCA	899
Db	1269	GAGAGAGGCGAGGAGAG-TGGAGGAGAGAGAGGCGGAGACAGAGAGAGGAAGAAGAGA	1327
Qy	900	GAAGAGGGCCCGCAGGAGAGAGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	956
Db	1328	GAGAGAGGCGAGGAGAGTGGAGGAGAGAGAGAGGCGGAGACAGAGAGAGGAGAGAGAG	1387
Qy	957	CCAGACAGCCGGAAG	1016
Db	1388	GCAGGAGGTAGAGGAG	1447
Qy	1017	CGGAGAGAGAGGAG	1076
Db	1448	GAGTAG	1507
Qy	1077	CCGAGCCACCCGAG	1136
Db	1508	TGGAGGAG--TG	1566
Qy	1137	AAAGCACACCCCAACGAG	1196
Db	1567	GAGGAG	1626
Qy	1197	CAG	
Db	1627	GAG	

Search completed: October 21, 2004, 07:12:04
Job time : 143 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 21, 2004, 05:53:10 ; Search time 1687 Seconds
(without alignments)

4082.128 Million cell updates/sec

Title: US-10-019-341-4_COPY_256_1599

Perfect score: 1344

Sequence: 1 aagacacagcagagacgcg.....gagaagagagaagacccgcg 1344

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3407233 seqs, 2561960514 residues

Total number of hits satisfying chosen parameters: 6814466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 19: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 20: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
C 1	94.8	7.1	42999	9	US-09-799-462A-17
C 2	94.8	7.1	42999	10	US-09-836-911A-17
C 3	94.8	7.1	42999	10	US-09-738-630-73
C 4	94.8	7.1	42999	13	US-10-125-767-17
C 5	94.8	7.1	42999	14	US-10-151-081-17
C 6	94.8	7.1	42999	15	US-10-287-313-17
C 7	94.8	7.1	42999	15	US-10-219-694-17
C 8	94.8	7.1	42999	17	US-10-808-689-17
C 9	88.8	6.6	2318	10	US-09-814-353-21388
C 10	88.4	6.6	1468	15	US-10-369-493-39710
C 11	87.6	6.5	1840	15	US-10-017-161-2359
C 12	87.6	6.5	1840	15	US-10-292-798-2003
C 13	83.2	6.2	42533	15	US-10-004-113-43

C 14	81.4	6.1	6668	15	US-10-311-455-1670
C 15	80.6	6.0	1343	13	US-10-001-843-70
C 16	80.6	6.0	3259	13	US-10-001-843-71
C 17	79.8	5.9	1880	14	US-10-198-846-12701
C 18	79.2	5.9	1547	18	US-10-369-493-39721
C 19	78.2	5.8	792	13	US-10-027-632-168583
C 20	78.2	5.8	792	13	US-10-027-632-168584
C 21	78.2	5.8	792	13	US-10-027-632-168585
C 22	78.2	5.8	792	13	US-10-027-632-168586
C 23	78.2	5.8	792	15	US-10-027-632-168583
C 24	78.2	5.8	792	15	US-10-027-632-168584
C 25	78.2	5.8	792	15	US-10-027-632-168585
C 26	78.2	5.8	792	15	US-10-027-632-168586
C 27	77.2	5.7	1041	10	US-09-814-353-14764
C 28	77	5.7	868	17	US-10-437-963-42302
C 29	77	5.7	876	13	US-10-001-857-17
C 30	77	5.7	1186	15	US-10-017-161-2113
C 31	77	5.7	1186	15	US-10-292-798-1759
C 32	77	5.7	1225	15	US-10-017-161-2239
C 33	77	5.7	1225	15	US-10-292-798-1945
C 34	76.2	5.7	5258	14	US-10-198-846-10433
C 35	76	5.7	89625	13	US-10-087-192-2020
C 36	76	5.7	90614	18	US-10-331-053-64
C 37	75.4	5.6	1267	13	US-10-001-843-45
C 38	75.4	5.6	1297	15	US-10-017-161-2193
C 39	75.4	5.6	1297	15	US-10-292-798-1839
C 40	75.2	5.6	1706	15	US-10-369-493-39722
C 41	74.8	5.6	803	17	US-10-437-963-72176
C 42	74.8	5.6	1704	9	US-09-834-975-817
C 43	74.2	5.5	627	17	US-10-021-323-9336
C 44	74.2	5.5	1750	15	US-10-017-161-2107
C 45	74.2	5.5	1750	15	US-10-292-798-1753

ALIGNMENTS

RESULT 1

US-09-799-462A-17/c

; Sequence 17, Application US/09799462A

; Patent No. US20020160970A1

; GENERAL INFORMATION:

; APPLICANT: Hadlaczky, Gyula

; TITLE OF INVENTION: ARTIFICIAL CHROMOSOMES, USES THEREOF

; AND METHODS FOR PREPARING ARTIFICIAL CHROMOSOMES

; NUMBER OF SEQUENCES: 34

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Heller Ehrman White & McAuliffe

; STREET: 4250 Executive Square, 7th Floor

; CITY: La Jolla

; STATE: CA

; COUNTRY: USA

; ZIP: 92037

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSEQ Version 1.5

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/799,462A

; FILING DATE: 10-Sep-2001

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/835,682

; FILING DATE: 10-APR-1997

; APPLICATION NUMBER: 08/695,191

; FILING DATE: 07-AUG-1996

; APPLICATION NUMBER: 08/682,080

; FILING DATE: 15-JUL-1996

; APPLICATION NUMBER: 08/629,822

; FILING DATE: 10-APR-1996

; ATTORNEY/AGENT INFORMATION:

QY	90	ACAGGAAGAGAGAGGGGCCAAAACGGGCCCCCGACAGAGAGAACACAGACCCCAAGCAG	149
Db	25982	ATAGGAACAGAGAGAGAGAGAGACAGAGAGACGACAGAAAAGAAAGAGAGAGGCGAG	25923
QY	150	GGGACCGCCACCGGCCAGGAGCAACCCAGCCCGCGGCACACCAAACGGGGACACAGA	209
Db	25922	ACAGCAGAGAAACACAGACAGACAGAGAAACACGCGCAGAAAGAGAGAGAGAGAG	25863
QY	210	GGGCCCGGACAAACCGAGAGAGAGAAACAACCCGGGCAAGGCCACCGGAAACAGCCGGAG	269
Db	25862	AGAAACAGACAGTAAAGGACAGAGAGAGAGACAGAGAGAGAAAGAGAAATGAAGACAGAA	25803
QY	270	CCGAG---CGCGCAGCAGGAGCGACCAAAGAAAGCAACAGAAACGCGCGCACCGCGC	326
Db	25802	GACAGACACAGTGAGAGAGGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGACAGACAGACAG	25743
QY	327	GACCAACGAGAGCAGAACCCCGAGCCGCCGAGAGCAGAGAGAGCGCACACACACACAG	386
Db	25742	GACAGAGAGAAAGACACAGACAGACAGAGAAAGACACAGACAGAGAAAGACAGAGAGAGA	25683
QY	387	AGGGCCCCGGAACGACGAAACCAAGAACCGGGGCAGACACACCGAGGAGGACCCAGC	446
Db	25682	AAGAGAGAAACAGGCAGAGAGAGAGAGCTAGCCAGAGAGAGAAACAGAAAGGTAGGGAGA	25623
QY	447	CAGGAGAACAGGAGAAAGCACCCCGAGCAGACAGAGAGGACCGAGAGAGGACCCAGCAGGAAG	506
Db	25622	CAGAGAGAGACACACAGACAGACAGATGACAGCAGAGAGAGAGAGTGAACACAGACAG	25563
QY	507	CCCCACGAGCGCCACACCCACGACCGCGAAGAGAAACCAAGAGAGGCCACAGGGCAGC	566
Db	25562	ACACAGTCAGAGAGACAAGGAGAGAGAGAGAGAGAGACAGAGACAGAGACAGAGACAG	25503
QY	567	CAAGCAAGCCGAGAAAGGCCCGCGAGCAAGAAACCCGCGCAACACCGCGCAGAGACAA	626
Db	25502	ACGACAGAAAGAGAGAGAGACAGACAGACAGAAACACAGAGACCGACAGAGAGAGA	25443
QY	627	AAGCAGAGCCAAAAGAGACGCAAAAGACCGCAAGACCGCCAGAGCCCCAACAGCCCAACC	686
Db	25442	CAGAGAAACAGAGAGAAAGAGAGAAAGACAGAGAGCGAGAGGCGAGAGAGAGAGAGAA	25383
QY	687	AGAAAGACACCGGACGAGAGGAAACCCACCAACAGCGCCGACACCGAGGCCACCGGCCG	746
Db	25382	CAGAAAGCGGAGAGAGACAGAGAGACACAGACAGATAGACAGCGCAGAAAGAGAGT	25323
QY	747	AGAGAGAAACACCCACACCGCCGAAAGCCACAAAGAACACCCCCCAACACAGAGAGAA	806
Db	25322	AAGACAGAGATAGGCACAGAGAGAGACAGAGAGACACAGAAAGAGAAAGAGAGAGGC	25263
QY	807	GGAGAAACACCGAAGCCAAAGGAAGAGACACAGCGGCAGACCGGGAGCAGCCCGGCC	866
Db	25262	AGACACAGAGAAAGGGGACAGACAGAGAAAGCAGAGACAGAGAGAG-----	25213
QY	867	GCCACAGAGACAGAGAAAGCAGGAGAGACACAGAAAGGGGACCGCAGGGAGAAAGGCC	926
Db	25212	-----AGAGAGAGAGAGAGAGAAACACACAGAAAGAGAGAGGACAGGGAGAGAG	25159
QY	927	AGCAGAAAGGAAGGACCGCGGAGGAAACCTAGAACCGCGAAAGAGAGAGCGCGAAC	986
Db	25158	AGACACAGACAGACGACGACAGGACAGAGAGAGGTAAAGCAAAAGATACACACAGAGA	25099
QY	987	GGCGACACACAGAACAGAACCGCAGGAAACGGAAGAGAGGAGGAGAAACACAAAC	1046
Db	25098	GGGAGAGACAGAGAGACAGAGACAGACAGACAGAGAGATGAACAGACAGACA	25039
QY	1047	AACCCAGGGGGGCAAAAGGGACCGAAAAACCCAGCCCCCAGAGAGAGGAGACGCCAG	1106
Db	25038	GAGAGAGACAGTGAGAAACACAGACAGAAAGAGAGGAGACAGAGAGAAACACAGAGGAG	24979
QY	1107	CACAAAAGGCGAACAAAGGGGAGGCCCAAGACACACCCCAACGAAACACAGAGGACA	1166
Db	24978	GNNAGAGAG-----AGAGAGAGAGAGAGACAGAAAGGGAGGGAGACAGAGAGA	24926

D6 25742 GACAGAGAGAAAGAGACAGACAGACAGACAG

Query Match	6.6%;	Score 88.4;	DB 15;	Length 1468;
Best Local Similarity	47.0%;	Fred. No. 5.4e-11;		
Matches 404;	Conservative 0;	Mismatches 447;	Indels 8;	Gaps 4;
QY	481	AGGCGGAGAGGACCGACGAGGAAGGCCCCACGACGCGCCACACCCACGACCCCGGAAGAA	540	
Db	210	AGAACAAAACGCAACACGCGGAATTCGCCCAAAACAAGTTTGCACAGAAAACGACGAGA	269	
QY	541	GAAACCAAGAAGGCCACAGGGCAGCCAAAGGAGCCGAGAAAGGGCCCGAGGGAAGAAGA	600	
Db	270	CACAAACACACACAGGGGAGAAAACAGAGGGGGGGGAAACACACACAGAAAAGAAGA	329	
QY	601	ACCGCGCAACACCGGGCAGAGACAAAAGACGAGCCAAAAGACGACAAAAGACCCGAAG	660	
Db	330	GACACAGAAAACACACACGACACGAGAAAAGACACACACACCAAGAGACCCGCCAAA	389	
QY	661	ACCGCGAGGCCACCAAGCCCAACCAAGAAAGAACACGGGACGAGAGGAACCCACCAAA	720	
Db	390	CCAGACACAAAAGACAAAGACAAACACACACGGCCCAACAGAAAACACACCGAAGCAGAAA	449	

```

RESULT 11
US-10-017-161-2359/c
; Sequence 2359, Application US/10017161
; Publication No. US20030143668A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 094335/0152
; CURRENT APPLICATION NUMBER: US/10/017,161
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP 2001/246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2430
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2359
; LENGTH: 1840
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)..(1840)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (201)..(1640)

```

1	FEATURE:	
2	NAME/KEY: modified_base	
3	LOCATION: (1)	
4	OTHER INFORMATION: a, t, c, g, unknown or other	
5	FEATURE:	
6	NAME/KEY: modified_base	
7	LOCATION: (12)	
8	OTHER INFORMATION: a, t, c, g, unknown or other	
9	FEATURE:	
10	NAME/KEY: modified_base	
11	LOCATION: (20)	
12	OTHER INFORMATION: a, t, c, g, unknown or other	
13	FEATURE:	
14	NAME/KEY: modified_base	
15	LOCATION: (29)	
16	OTHER INFORMATION: a, t, c, g, unknown or other	
17	FEATURE:	
18	NAME/KEY: modified_base	
19	LOCATION: (39)	
20	OTHER INFORMATION: a, t, c, g, unknown or other	
21	FEATURE:	
22	NAME/KEY: modified_base	
23	LOCATION: (48)	
24	OTHER INFORMATION: a, t, c, g, unknown or other	
25	FEATURE:	
26	NAME/KEY: modified_base	
27	LOCATION: (51)	
28	OTHER INFORMATION: a, t, c, g, unknown or other	
29	FEATURE:	
30	NAME/KEY: modified_base	
31	LOCATION: (61)	
32	OTHER INFORMATION: a, t, c, g, unknown or other	
33	FEATURE:	
34	NAME/KEY: modified_base	
35	LOCATION: (75)	
36	OTHER INFORMATION: a, t, c, g, unknown or other	
37	FEATURE:	
38	NAME/KEY: modified_base	
39	LOCATION: (77)	
40	OTHER INFORMATION: a, t, c, g, unknown or other	
41	FEATURE:	
42	NAME/KEY: modified_base	
43	LOCATION: (95) .. (96)	
44	OTHER INFORMATION: a, t, c, g, unknown or other	
45	FEATURE:	
46	NAME/KEY: modified_base	
47	LOCATION: (98) .. (99)	
48	OTHER INFORMATION: a, t, c, g, unknown or other	
49	FEATURE:	
50	NAME/KEY: modified_base	
51	LOCATION: (110) .. (112)	
52	OTHER INFORMATION: a, t, c, g, unknown or other	
53	FEATURE:	
54	NAME/KEY: modified_base	
55	LOCATION: (120)	
56	OTHER INFORMATION: a, t, c, g, unknown or other	
57	FEATURE:	
58	NAME/KEY: modified_base	
59	LOCATION: (126)	
60	OTHER INFORMATION: a, t, c, g, unknown or other	
61	FEATURE:	
62	NAME/KEY: modified_base	
63	LOCATION: (130)	
64	OTHER INFORMATION: a, t, c, g, unknown or other	
65	FEATURE:	
66	NAME/KEY: modified_base	
67	LOCATION: (134)	
68	OTHER INFORMATION: a, t, c, g, unknown or other	
69	FEATURE:	
70	NAME/KEY: modified_base	
71	LOCATION: (143)	
72	OTHER INFORMATION: a, t, c, g, unknown or other	
73	FEATURE:	

```

1 / NAME/KEY: modified_base
2 / LOCATION: (146)..(148)
3 / OTHER INFORMATION: a, t, c, g, unknown or other
4 / FEATURE:
5 / NAME/KEY: modified_base
6 / LOCATION: (150)
7 / OTHER INFORMATION: a, t, c, g, unknown or other
8 / FEATURE:
9 / NAME/KEY: modified_base
10 / LOCATION: (152)
11 / OTHER INFORMATION: a, t, c, g, unknown or other
12 / FEATURE:
13 / NAME/KEY: modified_base
14 / LOCATION: (155)..(156)
15 / OTHER INFORMATION: a, t, c, g, unknown or other
16 / FEATURE:
17 / NAME/KEY: modified_base
18 / LOCATION: (158)
19 / OTHER INFORMATION: a, t, c, g, unknown or other
20 / FEATURE:
21 / NAME/KEY: modified_base
22 / LOCATION: (161)..(162)
23 / OTHER INFORMATION: a, t, c, g, unknown or other
24 / FEATURE:
25 / NAME/KEY: modified_base
26 / LOCATION: (182)
27 / OTHER INFORMATION: a, t, c, g, unknown or other
28 / FEATURE:
29 / NAME/KEY: modified_base
30 / LOCATION: (186)
31 / OTHER INFORMATION: a, t, c, g, unknown or other
32 / FEATURE:
33 / NAME/KEY: modified_base
34 / LOCATION: (189)
35 / OTHER INFORMATION: a, t, c, g, unknown or other
36 / FEATURE:
37 / NAME/KEY: modified_base
38 / LOCATION: (192)..(194)
39 / OTHER INFORMATION: a, t, c, g, unknown or other
40 / FEATURE:
41 / NAME/KEY: modified_base
42 / LOCATION: (224)..(227)
43 / OTHER INFORMATION: a, t, c, g, unknown or other
44 / FEATURE:
45 / NAME/KEY: modified_base
46 / LOCATION: (235)
47 / OTHER INFORMATION: a, t, c, g, unknown or other
48 / FEATURE:
49 / NAME/KEY: modified_base
50 / LOCATION: (241)..(242)
51 / OTHER INFORMATION: a, t, c, g, unknown or other
52 / FEATURE:
53 / NAME/KEY: modified_base
54 / LOCATION: (249)
55 / OTHER INFORMATION: a, t, c, g, unknown or other
56 / FEATURE:
57 / NAME/KEY: modified_base
58 / LOCATION: (264)
59 / OTHER INFORMATION: a, t, c, g, unknown or other
60 / FEATURE:
61 / NAME/KEY: modified_base
62 / LOCATION: (266)
63 / OTHER INFORMATION: a, t, c, g, unknown or other
64 / FEATURE:
65 / NAME/KEY: modified_base
66 / LOCATION: (270)..(369)
67 / OTHER INFORMATION: a, t, c, g, unknown or other
68 / FEATURE:
69 / NAME/KEY: modified_base
70 / LOCATION: (409)
71 / OTHER INFORMATION: a, t, c, g, unknown or other
72 / FEATURE:
73 / NAME/KEY: modified_base

```

[illegible]

```

; NAME/KEY: modified base
; LOCATION: (61)..(61)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (75)..(75)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (77)..(77)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (95)..(96)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (98)..(99)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (110)..(112)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (120)..(120)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (126)..(126)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (130)..(130)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (134)..(134)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (143)..(143)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (146)..(148)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (150)..(150)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (152)..(152)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (155)..(156)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (158)..(158)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (161)..(162)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (182)..(182)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base

```

```

; LOCATION: (186)..(186)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (189)..(189)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (192)..(194)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (224)..(227)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (235)..(235)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (241)..(242)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (249)..(249)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (264)..(264)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (266)..(266)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (270)..(369)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (409)..(409)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (429)..(429)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (437)..(437)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (481)..(481)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (489)..(489)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (506)..(506)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (528)..(528)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (578)..(579)
; OTHER INFORMATION: a, t, c, g, unknown or other

```

Query Match 6.5%; Score 87.6; DB 15; Length 1840;
 Best Local Similarity 47.6%; Pred. No. 8.4e-11;

; Sequence 1670, Application US/10311455
 ; Publication No. US20030143606A1

RESULT 14

Query Match	6.1%	Score 81.4	DB 15	Length 6668
Best Local Similarity	43.8%	Pred. No. 2.4e-09		
- Matches	352	Conservative	0	Mismatches 451
			Indels	Gaps
			0	0
QY	536	AAGNAGAAACCAAGAGGCCACAGGCGCCACGCGAAGCGCCGAGGAGCGCGAGGAG	595	
Db	3351	AA		
QY	596	AAAGAACCGCGCAACACCGGCGAGACAAAAAGCAGAGCCAAAGACAGCAGCAAAAGAC	655	
Db	3291	ACAAAAAATAA	3232	
QY	656	CGAAGACCGCGACAGCGCCCAAAAGCCCAACCAAGAAAGACACGGGCGAGGAAACCCA	715	
Db	3231	AAAAACAAAAAATAAAAAAAAAAAAAAAAAACAAAAAATAAAAAAAAAAAAAAAAAAAA	3172	
QY	716	ACCAACAGCGCGAGACCGAGGCCACCGGGCGGAGAGGAGAACCCACACCGCGGAAGCCA	775	
Db	3171	AAACAAAAAACAAAAAATAA	3112	
QY	776	CAAAAAGACACCCCAACACACAGAGGAGGAGGAGAACACCCAGGAGGAGGAGG	835	
Db	3111	AAAAAATAACAAAAAATAA	3052	
QY	836	ACAACAGCGCGACAGCGGGGAGCAGCCCGCGCCACAGAAAGACAGAAAGACGAGGAG	895	
Db	3051	AAACAAAAACAAAAAATAA	2992	
QY	896	ACGAGAAAAGGACCGCAGGAGAAAGCCGACGAGAAAGGAGGACACCGCGAGGAAA	955	
Db	2991	AAACAAAAAATAAATAA	2932	
QY	956	GCCACACAAGCCGAAAGAGGACGCGGAAACCGGGCGACACAGAACACGACGGA	1015	
Db	2931	AAAAACAAAAAATAA	2872	
QY	1016	ACGGAAGAGAGGAGGAAGAACAAAAACAAACCCAGGGGGGCAAAAGGACCGCAAA	1075	
Db	2871	ACAAAAAATAAATAAAAAAATAAAAAAAAAAAAAAAAAACGAAAAAATAAAAAACAA	2812	
QY	1076	ACCGAGCCACCCGAGAGGAGACAGCCAAAGCACAAAAAGGCGCAACAAAGGGAGG	1135	
Db	2811	AAAAAATAAATAAAAAAATAAAAAAAAAAAAAAAAAAAAAAATAAAAAAATAA	2752	
QY	1136	AAAACACACCCACGAAAGACAGGAGCAGAAAGCGCGCGGAGAAAGAAAAAGGAGC	1195	

Sat Oct 23 17:44:56 2004

QY	943	ACGCGGAGGAAGCCAGACAAGCCGAAAGAGCAGCGGAAACGGCGCAACACAGAAACA	1002
Db	588	ACGGAGAGACATCGGCACATAGACAGAGGAGCCGCGGATACAAAGNAAAAGGAACA	647
QY	1003	-----AAGAAACGGCAGGAACGGAAAGAGAGGAGGAGAAACACAAACACCCAGGGGG	1057
Db	648	AACGGAAGATTGAGAAAATATGACGAGCGACGAAGCAACGACCGGAAACCCAGACCAGCGCG	707
QY	1058	GGCAAAAGGACCCGAATAACCCAGCCACCCGAGNAGGAGACAGCCAAAGGGG	1117
Db	708	AGAGGCAAGAAAGGAGCAGAGAAACAAAAGCCGACAGAGAAAGCCAGACGAAAGAC	767
QY	1118	CAACAAAGGGGAAGGGCCAAAGCACACCCCAACGAAAGACAGGGACAGAAAAGGGCC	1177
Db	768	AAGCACGAGCTACAGG--AGGAGCCAAAGATGAGAAAGAGAGAGAGAAAGAAACAC	825
QY	1178	CGAGAAAGAAAAGGAGCGCAGAGAAAAGGCCCGGCGAGGGCCAGCCAAAGGAG	1237
Db	826	GAAGCAACAGACGCGAACACAGGAGAGAGAGAGAAACAGAGGGAGACGAAAGAGCAG	885
QY	1238	AACCCCAAGGACGACCGGACGAGGCCCAACCAACACCAAGCCCAAGAAACAGAAAAGCC	1297
Db	886	AGGAGAGAGAAACGAAAGTAGGGAGCCAGAGAAACGAAACGAGAGTACAAACAGNA	945
QY	1298	CGCGGCACGAAACAGACCCACACACAGCAGAGAGAGAGAGA	1338
Db	946	CAGGGAAGAAAGAGACCAAGGACAAAGAGGAAAGAAACACA	986

Search completed: October 21, 2004, 09:14:16
Job time : 1694 secs

Result	Query				ID	Description
	No.	Score	Match	Length		
	1	120.2	8.9	2172	9	CL078016
	2	119.6	8.9	1812	9	AG430199
	3	118	8.8	997	9	CNS005TE
	4	117.4	8.7	1723	9	AG430464
	5	113.8	8.5	1389	9	AG386733
	6	113.8	8.5	1511	9	AG350114
	7	113.2	8.4	987	9	CNS00418
	8	112.8	8.4	1407	9	AG346588
	9	111.6	8.3	1397	9	AG346292
	10	111.2	8.3	1424	9	AG436078
	11	110.6	8.2	1154	4	B1252098
	12	109.8	8.2	1542	9	AG346153
	13	109	8.1	1318	9	AG383258
	14	108.4	8.1	1319	9	AG324175
	15	107.6	8.0	1923	9	AG360576
	16	107.4	8.0	1260	9	AG3278508
	17	107.4	8.0	1582	9	AG346604
	18	107.2	8.0	1418	2	BF683173
	19	106.6	7.9	1402	9	AG324055
	20	106.6	7.9	1519	9	AG279933
	21	106.6	7.9	1595	9	AG365509
	22	106.4	7.9	1189	4	B1523055
	23	106.2	7.9	1362	9	CL082692
	24	106.2	7.9	1795	9	CL043523

[illegible]

RESULT 2
AG430199/c

[illegible]

```
/tissue type="mixture of kidney and spleen"  
/clone_lib="MSMG01 Mouse Male BAC Library"
```

ORIGIN

[illegible]

RESULT 6

AG350114/c	AG350114	1511 bp	DNA	linear	GSS 02-JUN-2004
LOCUS	Mus musculus molossinus	DNA, clone:MSMG01-146K01.TJ			genomic survey
DEFINITION	sequence.				
ACCESSION	AG350114				
VERSION	AG350114.1	GI:47923424			
KEYWORDS	GSS.				
SOURCE	Mus musculus molossinus				
ORGANISM	Mus musculus molossinus				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				

REFERENCE	AUTHORS	TITLE	JOURNAL
REFERENCE	AUTHORS	TITLE	JOURNAL

- 1 Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.
BAC and Sequences of Library MSMg01
Unpublished
- 2 (bases 1 to 1511)
Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.
Direct Submission
Submitted (17-NOV-2003) Mashira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suenhiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: hattori@gsc.riken.jp, URL: <http://hgpc.gsc.riken.go.jp/>,
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
Clones are derived from the mouse BAC library MSMg01. For BAC
library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
Tsukuba Institute, Bio Resource Center,
The Institute of Physical and Chemical
Research (RIKEN) 3-1-1
Koyadai, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abe@rtc.riken.jp

COMMENT

Clones are derived from the mouse BAC library MSMg01. For BAC library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
Tsukuba Institute, Bio Resource Center,
The Institute of Physical and Chemical Research (RIKEN) 3-1-1
Koyadai, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abe@rtc.riken.jp

e-mail: PRIMERS

e-mail: abe@rtc.riken.jp
 PRIMERS
 Sequencing : TJ
 LIBRARY
 Vector : pBACe3.6
 R.Site 1 : EcoRI
 R.Site 2 : EcoRI

FEATURES

```

Accession 1 : E0011
Location/Qualifiers
1. .1511
/organism="Mus musculus molossinus"
/mol_type="genomic DNA"
/sub_species="molossinus"
/db_xref="taxon:57486"
/clone="MSMG01-146K01.TJ"

```

ORIGIN

Query Match	8.5%;	Score 113.8;	DB 9;	Length 1511;
Best Local Similarity	46.1%;	Pred. No. 7.4e-13;		
Matches 411;	Conservative 0;	Mismatches 467;	Indels 13;	Gaps 1;
Qy	448	AGAGAACAGGAGAACGACCGGGGAGCAGAGAGAGCGGAGAGGACCCAGCAGGAAGGC	507	
Db	1359	AGAAGAAGGAGGAGAGAAAAGAGAGAAATAAAGCAGAAGAGANAAGAAAGAGGAAGAA	1300	
Qy	508	CCCACGAGCGCCACACCCACACCGCGNAGAGAAACCAAGAGGCCACAGGGCAGCC	567	
Db	1299	AGCAAGNAGAAAGAAAGAGAGAGAGAAAGAGAAAAGAGAAAAGAGAGAAG	1240	
Qy	568	AAGGAACCGGAGAAAGGGCGGAGGAGAAAGAACCGCGCAACCAACGGCGGAGACACAAA	627	
Db	1239	GAGAAAAGAGAAAAAAGAGGAGAGAAAGANGAGAGAAAAAGAGAAAGAGAA	1180	
Qy	628	AGCAGAGCCAAAGAGAGCAGCAAAGACCGAGAGCCGACAGGCCCAAAAGCCCAACCA	687	
Db	1179	GNAGAGAAAAGAGAGAAAAGNAGAAAAAANAAGAAAAAGAAAAAGGAAAGAA	1120	
Qy	688	AGAAAGACACGGGACGAGAGGAAACCCCAACCAACAGGCCGAGCCGAGGCCACCGGCCGA	747	
Db	1119	AGAAAGAAAAGAGAAAAGAGAGAAAGAGAAAGAAAGAAAGGAAAGAGAGAGAGAGAG	1060	
Qy	748	GAGGAGAACACCCACACCGCGCGAAGCCACAAAAAGACCCGCCCAACACAGAGGAGAG	807	
Db	1059	AGAGAAAAAAAAGAGAAAAAGAGAGAAAAAAGAAAAAGAAAGAAAGAGAGAGAAAAA	1000	
Qy	808	GAGAACACGAGAGCCAAAGGAAGAGGACAAACAGCCGCGAGACGGGAGCAGCCGCCG	867	
Db	999	AAGAGGAAAAAGAAAAAGAAAGAGAGAGAGAAAAAAGAAAGAAAGAGAGAAAAAGAA	940	
Qy	868	CCACAGAGAGACAGAGAAAGCCAGAGAGACCGAGAAAGGGACCGCAGGGAGAAAGCCA	927	
Db	939	GAGNAAAGAGAGAGAGAGAAAAAAGAAAAAAGAAAAAAGAGAAAAAGAAAAAGAA	980	

PRIMERS

Sequencing : T7
 LIBRARY : pBACe3.6
 Vector : pBACe3.6
 R.Site 1 : EcoRI
 R.Site 2 : EcoRI

FEATURES

source
 1. .1407
 /organism="Mus musculus molossinus"
 /mol_type="genomic DNA"
 /sub_species="molossinus"
 /db_xref="taxon:57486"
 /clone="MSMg01-141M24.T7"
 /sex="male"
 /tissue type="mixture of kidney and spleen"
 /clone_lib="MSMg01 Mouse Male BAC Library"

ORIGIN

Query Match 8.4%; Score 112.8; DB 9; Length 1407;
 Best Local Similarity 45.9%; Pred. No. 1.2e-12;
 Matches 471; Conservative 0; Mismatches 536; Indels 19; Gaps 2;

QY 287 AAGCGACCAAGAAAGCAACAGAAACGCGCCGACCGGACCAACGAGGACGAGCAAGCC 346
 DB |||||
 QY 1378 AAGAGAGCAACAGGAAGAAATGGAAAGCAAGAAAGAAAGAAATTAAGGAGAAAGCAAAA 1319
 DB |||||
 QY 347 CCGAGCGCCCGAGAGCAGAGAGCGCACACACACACAGAGGGGCCCGGCGGAGCAGGA 406
 DB |||||
 QY 1318 GAAGAGAGGAGAGAAAGAAAGCAATTAAGAAACAAGCAAAAAAGAGAAAGAAAGA 1259
 DB |||||
 QY 407 ACTGAAACACAGGGGCGAGCAACCCGAGAGGAGCAGCCAGGAGGAGAAACAGGAGAGCAC 466
 DB |||||
 QY 1258 ANGGAAGCAAAAGAGAGAGAGAGAGAGAAACAAAGAAAGAAAGAAAGAAAGAAAGAAA 1199
 DB |||||
 QY 467 CGCGAGCAGAGAGAGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 526
 DB |||||
 QY 1198 GAAAGAGAGAGAAAGCA 1139
 DB |||||
 QY 527 ACAGCCCGGAG 586
 DB |||||
 QY 1138 AAAAAACAAGAAAAAAGAA 1079
 DB |||||
 QY 587 CGGAGCA 646
 DB |||||
 QY 1078 GGAAGAAA 1019
 DB |||||
 QY 647 GCAAG 706
 DB |||||
 QY 1018 AGAAGAA 962
 DB |||||
 QY 707 GGAAG 766
 DB |||||
 QY 961 AGAAGAA 902
 DB |||||
 QY 767 CGGAGCA 826
 DB |||||
 QY 901 AGAAGAA 858
 DB |||||
 QY 827 AGAAGAA 886
 DB |||||
 QY 857 AGAAGAA 798
 DB |||||
 QY 887 GCAGGAG 946
 DB |||||
 QY 797 AAGAA 738
 DB |||||
 QY 947 CGGAA 1006
 DB |||||
 QY 737 AGGAA 678
 DB |||||
 QY 1007 ACGGAG 1066
 DB |||||
 QY 677 AAGAA 618
 DB |||||

QY 1067 GACCGGAAACCCAGCCACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1126
 DB |||||
 QY 617 AGAAGAA 558
 DB |||||
 QY 1127 GGAAGGGCCAAAGCAGACCCCAACGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1186
 DB |||||
 QY 557 AAAGAA 498
 DB |||||
 QY 1187 AAAAGGAGCGCAG 1246
 DB |||||
 QY 497 GAAGAA 438
 DB |||||
 QY 1247 GGAAGGAGCGGAG 1306
 DB |||||
 QY 437 AAGAA 378
 DB |||||
 QY 1307 AACAGA 1312
 DB |||||
 QY 377 GGAAGA 372

RESULT 9
 AG346292/c
 LOCUS Mus musculus molossinus DNA, clone:MSMg01-141G14.TJ, genomic survey
 DEFINITION AG346292 1397 bp DNA linear GSS 02-JUN-2004
 ACCESSION AG346292
 VERSION AG346292.1 GI:47919602
 KEYWORDS GSS.
 SOURCE Mus musculus molossinus
 ORGANISM Mus musculus molossinus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1
 Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
 BAC end Sequences of Library MSMg01
 Unpublished
 2 (bases 1 to 1397)
 Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
 Direct Submission
 Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
 1-7-22 Suehiro-cho,Tsukuba, Ibaraki, Japan, URL: http://hgp.gsc.riken.go.jp/
 (E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/
 Tel:81-45-503-9111, Fax:81-45-503-9170)
 Clones are derived from the mouse BAC library MSMg01. For BAC
 library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
 Tsukuba Institute, Bio Resource Center,
 The Institute of Physical and Chemical Research (RIKEN) 3-1-1
 Koyadai, Tsukuba, 305-0074 Japan
 phone: 81-298-36-9189, fax: 81-298-36-9199
 e-mail: abe@rtc.riken.jp

COMMENT

PRIMERS
 Sequencing : T7
 LIBRARY : pBACe3.6
 Vector : EcoRI
 R.Site 1 : EcoRI
 R.Site 2 : EcoRI

FEATURES

source
 1. .1397
 /organism="Mus musculus molossinus"
 /mol_type="genomic DNA"
 /sub_species="molossinus"
 /db_xref="taxon:57486"
 /clone="MSMg01-141G14.TJ"
 /sex="male"
 /tissue type="mixture of kidney and spleen"
 /clone_lib="MSMg01 Mouse Male BAC Library"

ORIGIN

Query Match 8.3%; Score 111.6; DB 9; Length 1397;
 Best Local Similarity 45.9%; Pred. No. 2.1e-12;

ORIGIN	Query Match 8.1%; Score 108.4; DB 9; Length 1319;									
	Best Local Similarity 48.7%; Pred. No. 9.2e-12;									
	Matches 386; Conservative 0; Mismatches 387; Indels 19; Gaps 3;									
QY	535	GAAGAAGAAACCAAGAGGCCACAGGGCAGCCAGGAAGCCGAGAGAGGCCCGCGGAGGA	594							
Db	1113	GAACAAGAAAAAAG	1054							
QY	595	GAAGAAGCCGCAACACAGCGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	654							
Db	1053	AAAAAG	994							
QY	655	CCGAAGACCGCCAGAGCCCAACAAAGCCCAACAGAGAGAGAGAGAGAGAGAGAGAGAG	714							
Db	993	AAAAACAACAACAAG	934							
QY	715	AACCAACAGCGCGAGAGCCGAGCGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	774							
Db	933	AAGAAAG	874							
QY	775	ACAAAGACACACCCCAACACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	834							
Db	873	AAAAAG	814							
QY	835	GACACAGCGCGAGAGCGGGGAGCAGCCCGCGCCACAGAGAGAGAGAGAGAGAGAGAGAG	894							
Db	813	AAAAAG	766							
QY	895	GACCAAGAAAGGGACCCGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	954							
Db	765	GAG	710							
QY	955	AGCCAGAGAGCGGAG	1014							
Db	709	AG	650							
QY	1015	AACCGAAG	1074							
Db	649	CAAG	590							
QY	1075	AACCCAGCCACCCGAG	1134							
Db	589	AACAGGCAACAGGAG	530							
QY	1135	CAAAAGCACACCCCAACAG	1194							
Db	529	AAAAAG	471							
QY	1195	CGCAG	1254							
Db	470	--AGCAAG	413							
QY	1255	CCGAG	1314							
Db	412	AAAG	353							
QY	1315	ACACAG								
Db	352	AAAAAANNNAANA	341							
RESULT 15										
AG360576/c										
LOCUS										
DEFINITION				Mus musculus molossinus DNA, clone:MSMg01-161B05.TJ, genomic survey						
				sequence.						
ACCESSION				AG360576						
VERSION				AG360576.1						
KEYWORDS				GSS.						
SOURCE				Mus musculus molossinus						
ORGANISM				Mus musculus molossinus						

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
BAC end Sequences of Library MSMg01
Unpublished
2 (bases 1 to 1923)
Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
Direct Submission
Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC); Japan
1-7-22 Suehiro-chou,Tsuri-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@psc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
Clones are derived from the mouse BAC library MSMg01. For BAC
library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
Tsukuba Institute, Bio Resource Center,
The Institute of Physical and Chemical Research (RIKEN) 3-1-1
Koyadai, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abe@rtc.riken.jp
PRIMERS
Sequencing : TJ
LIBRARY : pBACE3.6
Vector : EcoRI
R.Site 1 : EcoRI
R.Site 2 : EcoRI
Location/Qualifiers
1.1923
/organism="Mus musculus molossinus"
/mol_type="genomic DNA"
/sub_species="molossinus"
/db_xref="taxon:57486"
/clone="MSMg01-161B05.TJ"
/sex="male"
/tissue type="mixture of kidney and spleen"
/clone_lib="MSMg01 Mouse Male BAC Library"

Query Match 8.0%; Score 107.6; DB 9; Length 1923;
Best Local Similarity 45.4%; Pred. No. 1.4e-11;
Matches 352; Conservative 0; Mismatches 422; Indels 1; Gaps 1;

QY 564 AGCCAGGAGAGCGGAG 623
Db 1921 AGAAAG 1862
QY 624 AAAAAAG 683
Db 1861 AAAAAAG 1802
QY 684 ACCAAG 743
Db 1801 AAAAAAG 1742
QY 744 CCGAG 803
Db 1741 AG 1682
QY 804 GAAG 863
Db 1681 GAG 1622
QY 864 GCCGCCAG 923
Db 1621 AAAAAAG 1562
QY 924 GCCAG 983
Db 1561 GAAAAAG 1502
QY 984 AACGGGAG 1043

Db 1501 AAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAA-AAAA 1443
QY 1044 AACAAACCCAGGGGGCAAAAGGACCGAAAAACCCAGCCACCCGGAAGAGACAGCC 1103
Db 1442 AAAAAGAGAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 1383
QY 1104 AAGCACAAAAAGGGCAACAAAGGGAGGGCCAAAAAGCACACCCCAACGAAAGACAGGG 1163
Db 1382 AAAAAAAGAAAGAGGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 1323
QY 1164 AAGAAAAAGCGCCCGAGAAAGAAAAAGGAGCGCAGAGAAAAAGGCCCCAGGGCGAGGG 1223
Db 1322 AGAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA 1263
QY 1224 CCAAGCCAAAGGAGAACCCCAAGAGAACGACCGGACCGAGGCCCCCAACCAAGCCCA 1283
Db 1262 AAAAAAGAAAGAAAGAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAA 1203
QY 1284 GAAACAGAAAAAGCCCGCGCACCAAGAGAACACACACAGCAGAGAGAGAGAGAA 1338
Db 1202 AGAAAAAAGAGAAAAAGAGAGAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAA 1148

Search completed: October 21, 2004, 07:09:42
Job time : 4606 secs

This Page Blank (uspto)